What is the weather usually like?

Climate averages and normals: Long term statistics of weather data

Calculation of normals 1
30 year standard normal 1
Information on the CD 2
Missing data 2
Extremes 2
Length of record and missing data 2
Limited elements 3
Statistics used 3
Means for 9 am and 3 pm and Daylight Saving Time 3
Site numbers and rainfall districts 4

Calculation of normals

The calculation of climate normals or averages over an extended period of time, provides an indication of the typical climatic values for any particular location. Means and other values from shorter periods of time (eg. 1 day, 1 month or 1 year) can be compared to the relevant normal value to determine how unusual that period was ie. whether it was above or below average, and by how much.

There are a number of different ways of defining a normal, such as simply calculating a long-term average using all years of available record. This method has the advantage that a normal can be calculated for any climate record, no matter when it starts or ends. However, due to the fact that the climate records from different sites often span different time periods, intercomparisons between their normals may not be possible.

30 year standard normal

Alternatively, a standard reference period can be defined over which the statistics are calculated. Typically this is 30 years, the current international standard reference period being 1961 to 1990. This method allows normals calculated from different climate records to be compared. For example, in a map of average monthly rainfall all data points must be calculated with respect to the same period of normal, otherwise some data points might have been calculated during relatively wet times, others during relatively dry periods, and the map will not provide a true indication of the rainfall distribution throughout the region.

A 30 year period provides statistics considered stable enough not to be influenced by particularly unusual years and is also more relevant to people, being approximately equal to many people's length of memory. For example, if a temperature record has steadily warmed over the past 100 years a normal calculated using all years of record would be less meaningful than one calculated over a recent 30 year period, one to which people might actually be able to relate. However, this approach has the disadvantage that many sites may not have sufficient data during the defined period of normal to calculate a reliable value.

As an example: DARWIN AIRPORT

CLIMATE NORMALS _ 1961 and 1990

	JAN	***JUL***	***ANN***
Mean Daily Max Temp (deg C)	31.8	30.5	32.0
Mean Rainfall (mm)	436.6	1.0	1705.1
Median (Decile 5) Rainfall (mm)	425.7	0.0	1685.1

CLIMATE AVERAGES _ 59 years of record to 2000

	JAN	***JUL***	***ANN***
Mean Daily Max Temp (deg C	31.7	30.4	31.9
Mean Rainfall (mm)	428.5	1.0	1714.9
Median (Decile 5) Rainfall (mm)	403.7	0.0	1686.8

Information on the CD

On this CD we have provided both the long term climate statistics using all available years of record and the climate normals for the period 1961-1990. There are some differences immediately obvious: there are some 1100 sites in the all-available years dataset, and only some 480 sites in the standard period dataset. Further we have not included the extreme values, such as highest maximum temperature, in the standard period normals as it is usually felt that a much longer period is required to provide reliable values for these. However this test has not been applied to the extreme values in the all-available years dataset, so care should be used.

When selecting sites for the standard 30 year normal, we still included those with at least 20 years of maximum air temperature data as there were only about 140 sites with a virtually complete record. This is not as rigid as the specifications laid down by the World Meteorological Organisation.

We have also included the rainfall statistics only, for sites with more than 20 years of monthly rainfall data - about 12,000 sites.

Missing data

Incomplete ('short') months

There are many sites which have not recorded a full month of daily values; in some cases these 'short' months may only have a few days from which the mean daily maximum temperature for example can be calculated. Because of this for most elements the output uses only months which have more than 20 days of observations. Sites which do not report on weekends will usually be included, but their percentage completeness (given at the end of the row) will be lower (about 70%) as compared to a complete unbroken record.

Extremes

Extremes in 'short' months

All data were used for the Highest maximum or lowest minimum temperatures, or for the maximum wind gust as these values are not biassed by 'short' months to the same extent.

Early and very recent extremes

The temperature extremes given in this report are those values which are in the computer archive. Some sites may have had more extreme values in the 1800s or the early 1900s, which have not yet been computer entered. Extremes which have occurred within the last two to three months of this report, may also not have been entered into the computer archive.

Highest daily rainfall

The highest daily rainfall given here is the highest value which has been recorded. Many sites report accumulated falls at the end of a weekend or holiday, and such falls may conceal higher daily amounts than are shown, although this is impossible to identify due to lack of information. These accumulated falls were not included in the calculation of the highest daily total.

Length of record and missing data

The output indicates the length of the record and how complete that record might be. Together these supply a rough indication of what data are available for the element in question.

! The field '*years of record'* is simply the number of months used divided by 12, and does **not** mean calendar or complete years except for the percentile values. It gives the rough amount of data used between the first date of occurrence and the last data of occurrence *of the element in question*.

! The field '*percentage complete'* gives an estimate of how complete the data are for that element, where a record with no missing values between the first and last dates would be 100% complete. Sites with missing data will be less than 100%.

The *'percentage complete'* and the '*years of record'* for one element are not necessarily related to those for another element. There will often be far more data available for rainfall than there will be for other elements. Similarly there are normally far more years of mean maximum and minimum temperature data, than there are of eg the number of days at or above 40 ^{o}C.

The rainfall percentile figures were derived from 'complete' years that is years with no missing monthly totals. Thus the number of years will often be less than for the monthly mean.

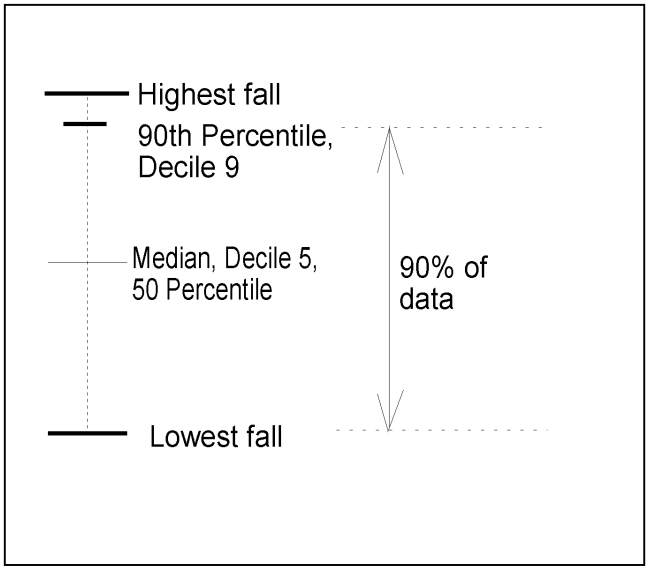
Limited elements

In many cases there will be no data for a particular month or element. The rows for sunshine duration, mean wind speed, maximum wind gust or evaporation will often not be included as these elements are only recorded by a restricted number of sites.

Statistics used

Median and percentiles

To calculate deciles (percentiles), we divide the ranked dataset into ten parts. The median is simply that value which marks the level dividing this dataset in half.



For example 50% of Januaries will have a total rainfall above the January median and 50% will have a total below. The median is also known as the 5th decile, decile 5 and the 50th percentile _ they are all the same thing.

Decile 9 or the 90th percentile for January, signifies that 90% of January totals will be below this figure. In other words there is a 90% chance of a January rainfall being below decile 9 (90th percentile), a 10% chance of it being above decile 9. Similarly there is a 10% probability of it being below decile 1 (10th percentile).

To get the annual percentile or decile value, you do *not* sum the percentiles for the 12 individual months - you must calculate it separately. However it is possible for the two values to be the same by chance.

Average rainfall

Both mean and median (decile 5, percentile 50) rainfall are included in the tables, although median is the preferred measure of 'average' rainfall from the meteorological point of view. This is because of the extremely high variability of daily rainfall _ one very large fall will affect the arithmetic mean more than it should, but will have less effect on the median. The median is therefore usually considered the more reliable indicator. It is commonly lower than the mean.

Statistics and length of record

Within the relevant period, and subject to the previous selection conditions, all observations for a site that have been quality controlled were used, regardless of how many years of data there were.

Users should be aware that in Australia, rainfall especially is extremely variable and a period of less than 30 years of rainfall data may not produce reliable statistics. As a comparison some 5_10 years of temperature data will provide a reasonable estimate of the mean, (although not of the extremes). Such information should be used with caution.

Means for 9 am and 3 pm and Daylight Saving Time

Due to the effect of Daylight Saving, values at 9am and 3pm are only nominal for most Australian sites. Daylight Saving has been used in some, but not all, states of Australia, since about 1973. The changeover occurs almost always in October and March but the exact dates vary from state to state and year to year. The averages for a particular hour are hence generally a combination of 8am and 9am values, or of 2pm and 3 pm values. The effect is especially marked at 9 am when the air temperature is often rising quite sharply. See the Bureau's web site at: http://www.bom.gov.au/climate/averages/tables/daysavtm.shtml for more details.

Site numbers and rainfall districts

There is an explanation on the web at:
http://www.bom.gov.au/climate/averages/tables/site_num.shtml for the site numbers and the rainfall districts used by the Bureau's systems.