



WRL reference	M05 D01
Module	M05 Natural Resource Use and Sustainability
Data set	D01 Identifying overfishing on Indonesian coral reefs

Data collection methods:

During July and August of 2005 and 2011, fish fence catches were monitored by fishery scientists on the island of Kaledupa in the Wakatobi Marine National Park, Indonesia. Fishermen emptied their fences once per day as usual, and returned their catches to a central location where the monitoring team was able to record details of the catch. The total weight of the catch from each fence was measured, and the length of each individual fish was measured. The weight was used to represent catch per unit effort (CPUE, kg of fish per fence per day), and the length data used to calculate which individuals were adults and which were juveniles using published length of maturity for each species. In both study years, the total number of fences around Kaledupa Island was counted using a boat census, and the length of each fence measured to give an average fence length. The number of fences and the length of fence were used to represent the amount of effort being put into fish fence use around Kaledupa over time.

Analysis methods:

TO BEGIN WITH OPEN THE FILE "M05 D01 BLANK.xls" (located in the folder 4. Data for tasks)

n.b. the instructions below are for Microsoft's Excel spreadsheet (2007) for the PC and the detailed procedures may will differ for other platforms e.g. Macs.

In this spreadsheet you will see the raw data for catch per unit effort (CPUE), the proportion of juveniles, number of fish fences, and fence length. This data is repeated for both 2005 and 2011. Catch data is shown for 12 individual catches, while effort data is shown as mean (average) values for the whole island.

The first thing you need to do is summarise the catch data by completing some simple analyses. When you are working with a number of samples, the first step is always to calculate a mean (average) for each data set. Calculating a mean value in Excel is easy, and you can use the following formula (using 2005 CPUE as an example):

=AVERAGE(B3:B14)

However, it is not enough to just show the mean values, as you also need to account for the amount of variation in the data (e.g. the mean for 2011 may be very different to 2005, but the difference may only be real if the variation in each data set isn't too large). Standard deviation is a common way to show variation in data which many of you will be familiar with. However, in research a more common method is to use standard error, which takes into consideration how many replicates have been completed (the more replicates the more confident you can be in your mean value). To calculate standard error you simply divide the standard deviation by the square root of the number of replicates (here we have 12 replicates for each data set, because we have monitored 12 individual catches). Using the 2005 CPUE data (column B) as an example, you can calculate standard error by inserting the following formula into cell B17:





=STDEV(B3:B14)/SQRT(12)

You will need to repeat both the mean and standard error formulas for both 2005 and 2011 data for CPUE and proportion of juveniles, remembering to modify the formula each time (e.g. use the letter C for CPUE 2011 instead of the letter B because you are using data in the C column).

Looking at the standard error results, we can see that the lowest level of variation in the data was for CPUE in 2011 (0.76), and the highest was for the proportion of juveniles in 2011 (5.88). This could technically indicate that the mean value for 2011 CPUE could be treated with the most confidence. However, all of the standard error values are reasonably low compared to the means themselves, which is a positive sign that the mean values are all a good representation of what is happening in the fishery.

Now you have your data summarised it is much easier to describe, and also ready to be used to create a suitable figure (graph) later on.

Next you need to find out whether there is actually a real difference between the 2005 and 2011 data sets, and to do this a statistical test is needed. The easiest statistical test to look for a significant difference between two similar data sets (i.e. to look for a difference between two mean values) is a t-test. Here we will use an un-paired t-test because, although we are studying the same broad population at two different time points, we have not used fish fences positioned in exactly the same place, and are not measuring the same exact fish, and so we cannot truly say that our samples are paired. If in doubt, it is always better to use an un-paired test, as it is a more conservative test and therefore is less likely to tell you that there is a significant difference if there isn't.

Luckily Excel makes un-paired t-tests easy, and you can use a formula to do the calculation for you. An "Un-paired t-Test probability" box has already been prepared, with blank boxes for the results comparing the 2005 and 2011 data for both CPUE and proportion of juveniles. The formula to use (using CPUE as an example) is as follows:

=TTEST(B3:B14,C3:C14,2,2)

In this formula, the two sets of cells (B3:B14 and C3:C14) are the data sets being compared. The number 2 tells Excel you are using two tails (a normal distribution which is the most common), and the number 1 tells Excel this data is paired.

Now repeat the above formula but modifying it for the proportion of juveniles data sets.

TO TRY CALCULATING T-TEST RESULTS BY HAND, TAKE A LOOK AT "MANUAL T-TEST INSTRUCTIONS" IN THE DATA FOLDER

The results from these formulae show the probability that there is NO significant difference between the two years. In other words, the smaller this value is (known as a p-value, or probability value), the higher the chance that there is a significant difference between the two year. In research, we generally work to the principle that 95% or above confidence is significant, which translates to a p-value of 0.05 or below (i.e. a 5% chance that there isn't a significant difference). Both of the t-tests used here should give values lower than 0.05, allowing you to confidently say there is a significant difference between the data in 2005 and 2011 for both CPUE and the proportion of juveniles caught.

OPEN THE FILE "M05 D01 CALCULATED.xls" TO SEE WHAT THE ANALYSED DATA SHOULD LOOK LIKE





The final stage in this exercise is to use a suitable figure (graph) to illustrate the patterns in the data set. The best type of figure to use for this kind of data (i.e. two individual data sets collected in different years) is a bar graph. Because CPUE and the proportion of juveniles use completely different units and aren't comparable measurements, it is easiest to use separate figures for each data set. You could include a second y-axis on a single figure, but this is complicated and can easily confuse things.

To create a bar graph in excel, select the data you want to show (e.g. the mean values for CPUE), select 'Insert', and then click on 'Column'. This should create a basic bar graph showing the mean data. You will still need to add suitable axis titles (with units), a title, add standard error bars, and generally make the graph more presentable. The options you will need to make these changes can be found by clicking on the graph and selecting 'Layout'.

OPEN THE FILE "M05 D01 COMPLETED.xls" TO SEE OUR SUGGESTED FIGURES

