Exercise: Statistics

> **AIM:** To introduce you to the statistics functions in the genutil package.

Issues covered:

- The genutil package
- Correlation function
- Weights in statistics
- The 'squeeze' keyword

Instructions

1. Open the file "~/my_cdat_files/data/unknown.xml".

2. Extract the 'lsp' (large scale precipitation) and 'cp' (convective precipitation) variables.

3. Calculate the correlation between the variables.

   You will need to import the "genutil" package and use "genutil.statistics.correlation()".

4. Look at the shape of the resulting variable. Default axis is the first one found.

5. Now calculate the correlation across all dimensions.

6. Look at the shape and value of the result.

7. Write a "for" loop to print the correlation of the entire surface fields at each time step (12 in all).

8. Now you want to get the standard deviation of the 'lsp' variable over land only, read in the 'lsm' variable (land sea mask) from the file "~/my_cdat_files/data/lsm.nc".

9. Extract the first time step of lsp and print the standard deviation of the lsp variable across y and x axes.

10. Print the same thing using the lsm mask as the weights array (1=land, 0=ocean). Check the result is different from (9). You'll need to 'squeeze' lsm.