# Exercise: Aggregating data

**AIM:** To introduce the cdscan utility, CDML format and accessing data in CDMS datasets.

**Issues covered:**

- cdscan options
- CDML format
- Adding attributes

## Instructions

1. Go to the ~/my_cdat_files/data/ directory and create a file containing the full paths of all files matching the glob pattern "*_2000??.nc" .

***If you do this in python you might need "glob.glob().***

2. Run the "cdscan" utility on those files to create the CDML file: "rain_dataset_2000.xml".
3. Open the file in CDAT as you would any file and extract the 'lsp' (large scale precipitation) variable.
4. Take a time slice from the 6$^{th}$ to the 8$^{th}$ time steps.

***Remember you can use a "selector" or a "slice" object if you like.***

5. Have a look inside the CDML and observe the template has been inserted. Also note the axes are listed and finally the variables.
6. Now you decide you really need to add some metadata. Use cdscan to add the following global and variable attributes:
   Global: source="This data is top secret"
   Global: comment="It will self-destruct in 1 se-..."
   lsp: long_name="Large Scale Precipitation Monthly mean"
   cp: long_name="Convective Precipitation Monthly mean"
7. Now run CDAT again, extract both variables, sum them to get the total precipitation. Print and plot to see variable.
8. Save your new variable to:
   "~/my_cdat_files/output/tp_2000_means.nc".