

Data preparation for Question 1

The data file, `tornado.csv`, has one line per tornado. What is required for analysis is a file with one line per month for each month from 2000 to 2004. You will need to use `factor()` to attach labels to the integer coded variables and then apply `table()` and `tapply()` to create tables of counts. These tables will then be strung out into vectors to create variables with number of deaths, number of tornados and number of killer tornados, as follows:

1. Convert the year and month variables to factors, with appropriate labels
2. Use the `table()` command to create a table of counts of tornados in each month of each year.
`table(month,year)`
3. Use the `tapply()` command to sum the numbers of deaths by year and month
`tapply(deaths,list(month,year),sum)`
4. Create a dummy variable to represent killer tornados. Use `tapply()` as in the previous step to tabulate the number of killer tornados by year and month
5. Convert the previous tables into column vectors by using `as.vector()`
6. Month and year variables for the tabulated data are most conveniently obtained by apply the `row()` and `col()` commands to any of the matrices created in steps 2 to 4. Be sure to convert the results to factors.
7. For tidyness, put all the variables into a `data.frame`.