Statistics 13   **Instructor: Nicolas Christou**
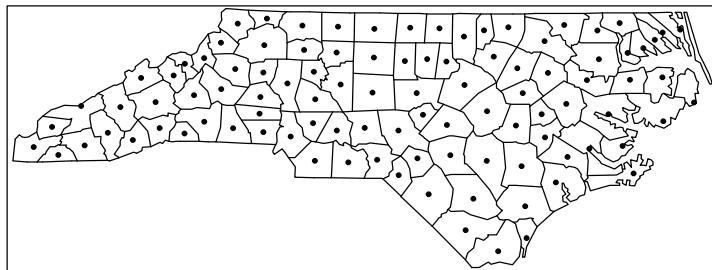
North Carolina 1974-1978 and 1979-84 SIDS data

Sudden Infant Death Syndrome (SIDS) data are collected for each of the 100 counties of the state of North Carolina and the total number is reported for the years 1974-1978 and 1979-84. The data points are assumed to be the county seats (see figure below).

Sudden Infant Death Syndrome (SIDS), is the sudden death of any infant up to 12 months old, that is not foreseen by the family history and where the postmortem examination does not give any logical explanation as to why the infant has died. SIDS is, or at least it was until very recently, the main cause of postneonatal death. It leads to about 7000 deaths per year in the United States - it is responsible for the deaths of about two infants per 1000 live births. See, Cressie, N. (1993). *Statistics for Spatial Data.* John Wiley, New York, pp. 244-248, 386-389.

You can access the data in `R` as follows:

```
a <- read.table("http://www.stat.ucla.edu/~nchristo/statistics13/nc_sids_data.txt", header=TRUE)
```

In addition, you can access the coordinates (longitude/latitude) of each county seat of North Carolina here:

```
b <- read.table("http://www.stat.ucla.edu/~nchristo/statistics13/nc_seats.txt", header=TRUE)
```

Work to do:
Analyze these data set in any way you want. You can use among other things the following: summary statistics, plots (boxplots, histograms, scatterplots, bubble plots), maps, regression analysis, variable transformations, etc. See the variable description on the next page. You are also allowed to use any external data that you think may help your analysis.

Description of the variables in data `nc_sids_data.txt`:

`name`	Name of county.
`i`	Number.
`x`	Distance (easting) of an arbitrary chosen origin (SW of North Carolina).
`y`	Distance (northing) of an arbitrary chosen origin (SW of North Carolina).
`births1`	1974-1978 births.
`sids1`	1974-1978 SIDS.
`birthsnw1`	1974-1978 nonwhite births.
`births2`	1979-1984 births.
`sids2`	1979-1984 SIDS.
`birthsnw2`	1979-1984 nonwhite births.

Description of the variables in data `nc_seats.txt`:

`name`	Name of county.
`lon`	Longitude of the county seat.
`lat`	Latitude of the county seat.

Please work with the following variables:

```
#Get number of SIDS per 1000 births:
Q1 <- 1000*(a$sids1+1)/a$births1     #1974-1978 data
Q2 <- 1000*(a$sids2+1)/a$births2     #1979-1984 data
```