# Data

The atmos data set resides in the nasaweather package of the R programming language. It contains a collection of atmospheric variables measured between 1995 and 2000 on a grid of 576 coordinates in the western hemisphere. The data set comes from the 2006 ASA Data Expo.

Some of the variables in the atmos data set are:

temp - The mean monthly air temperature near the surface of the Earth (measured in degrees kelvin (K))

pressure - The mean monthly air pressure at the surface of the Earth (measured in millibars (mb))

ozone - The mean monthly abundance of atmospheric ozone (measured in Dobson units (DU))

You can convert the temperature unit from Kelvin to Celsius with the formula

*<!-- Insert the conversion formula here -->*

And you can convert the result to Fahrenheit with the formula

$$ fahrenheit = celsius \times \frac{9}{5} + 32 $$

## Cleaning

To analyze this data, we will use the following R packages:

*<!-- Insert code chunk 1 here -->*

*<!-- Insert code chunk 2 here -->*

For the remainder of the report, we will look only at data from the year *<!—code chunk 3: Insert inline code to reference year -->* . We aggregate our data by location, using the \*R\* code below.

*<!-- Insert code chunk 4 here -->*

where the `year` object equals *<!-- code chunk 5: Insert inline code to reference year -->*.

## Ozone and temperature

Is the relationship between ozone and temperature useful for understanding fluctuations in ozone? A scatterplot of the variables shows a strong, but unusual relationship.

*<!-- Insert code chunk 6 here -->*

We suspect that group level effects are caused by environmental conditions that vary by locale. To test this idea, we sort each data point into one of four geographic regions:

*<!-- Insert code chunk 7 here -->*

### Model

We suggest that ozone is highly correlated with temperature, but that a different relationship exists for each geographic region. We capture this relationship with a second order linear model of the form

$$ ozone = \alpha + \beta\_{1} temperature + \sum\_{locales} \beta\_{i} locale\_{i} + \sum\_{locales} \beta\_{j} interaction\_{j} + \epsilon$$

This yields the following coefficients and relationships.

*<!-- Insert code chunk 8 here -->*

*<!-- Insert code chunk 9 here -->*

### Diagnostics

An anova test suggests that both locale and the interaction effect of locale and temperature are useful for predicting ozone (i.e., the p-value that compares the full model to the reduced models is statistically significant).

*<!-- Insert code chunk 10 here -->*