STEP 1

# About the Challenge Expo

The Annual Data Challenge Expo is jointly sponsored by three American Statistical Association (ASA) Sections – Statistical Computing, Statistical Graphics, and Government Statistics.

# Data

*<!-- Turn the line that begins with "Data" into a second level header.-->*

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 ASA Data Expo.

*<!-- Change the words atmos and nasaweather into a monospaced font suitable for code snippets.-->*

*<!-- Make the letter R italicized.-->*

*<!-- Change "2006 ASA Data Expo" to a link that points to*

*https://community.amstat.org/dataexpo/home.-->*

STEP 2

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))

*<!-- Turn the above text into a bulleted list with 3 bullets: temp, pressure, ozone.-->*

*<!-- Make temp, pressure, ozone bold at the start of each entry.-->*

*<!-- Make K, mb, and DU italicized at the end of each entry.-->*

STEP 3

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

*<!-- Insert the conversion formula celsius* = *kelvin* – 273.15 *here and display it on a separate row -->*

And you can convert the result to Fahrenheit with the formula

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

STEP 4

## Preparing the Data

To analyze this data, we will use the following R packages: library(nasaweather) and library(tidyverse)

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

*<!-- Set message=FALSE so messages are not produced in your report when you load the packages-->*

For the remainder of the report, we will look only at data from the year 1995. We aggregate our data by location, using the \*R\* code below.

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

STEP 5

## 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 3 here -->*

*<!-- Set the code chunk options: echo = FALSE, fig.height = 4, fig.width = 5, fig.align = "center"-->*

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 4 here -->*

# Conclusions

We suggest that ozone is highly correlated with temperature, but that a different relationship exists for each geographic region.