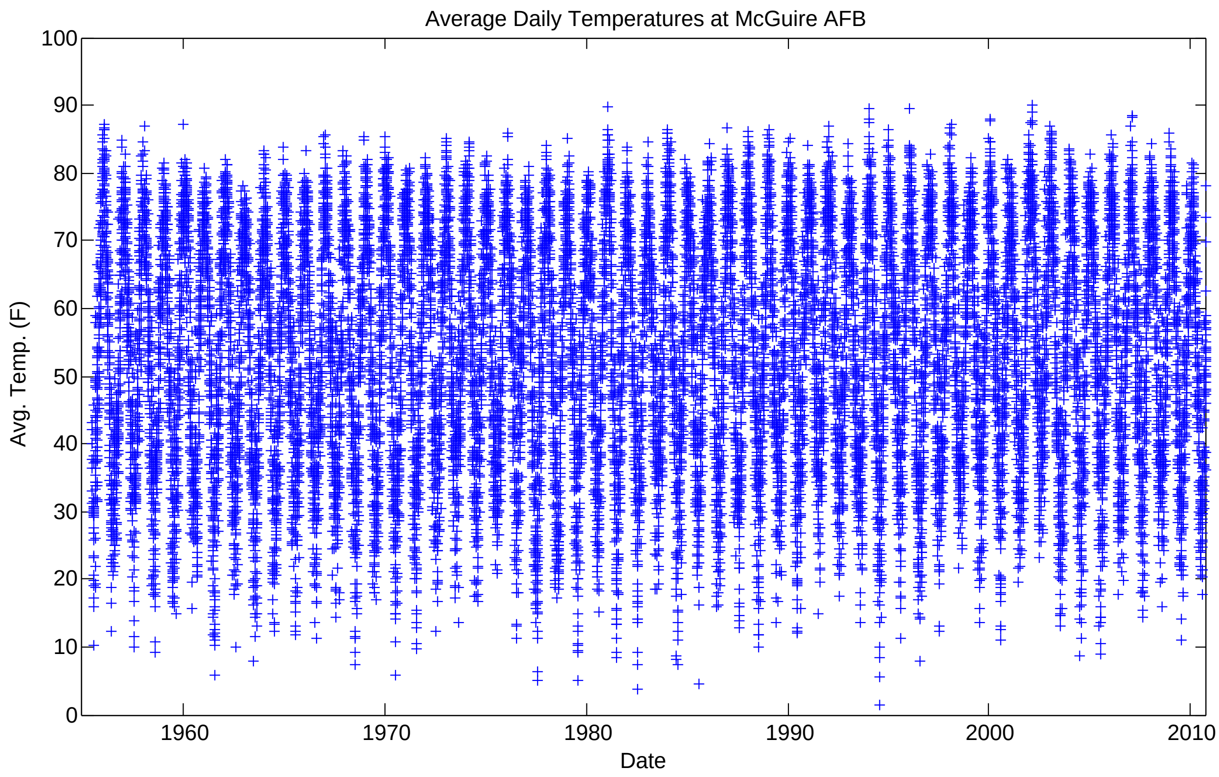
Two AMPL Problems for Homework 3

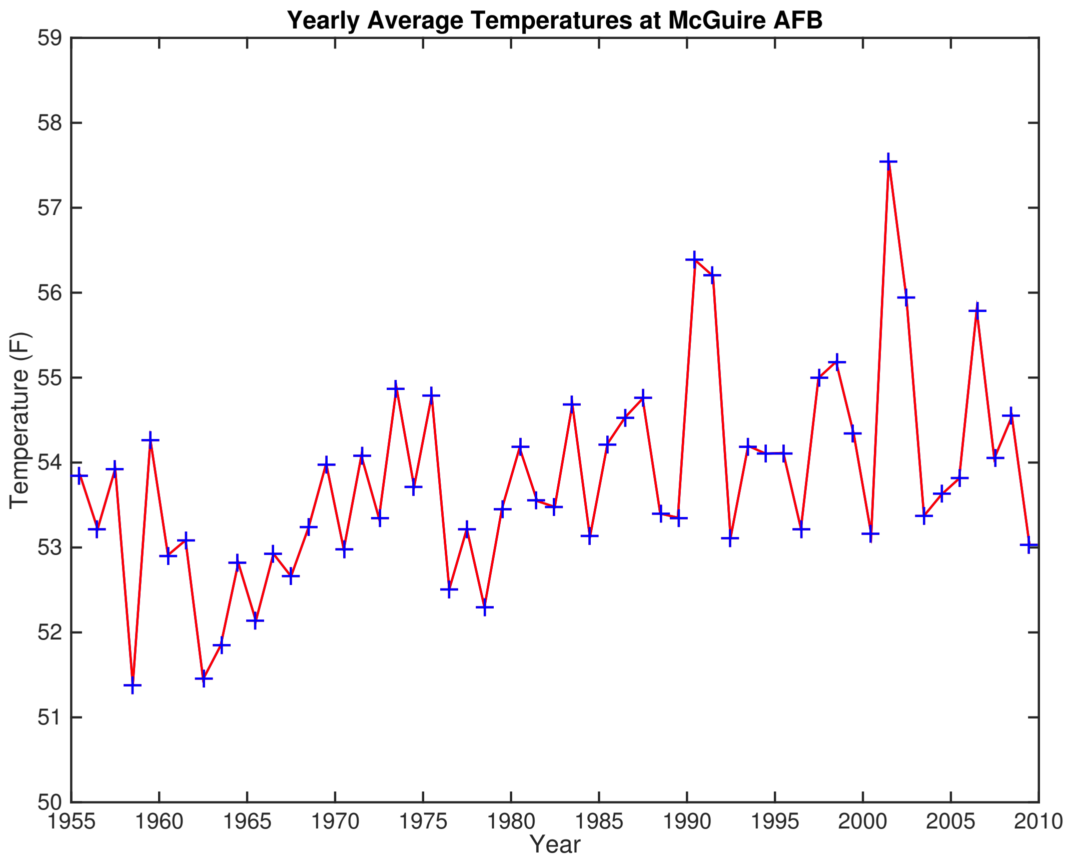
A few years ago I downloaded some data from the National Oceanic and Atmospheric Administration (NOAA):

`https://www.ncdc.noaa.gov/cdo-web/datatools`

The data is 55 years of daily average temperatures at McGuire Air Force Base, here in balmy NJ:



Using this raw data, I computed 55 yearly averages:



Here's the data in tabular form:

1955	53.85	1965	52.13	1975	54.78	1985	54.22	1995	54.11	2005	53.81
1956	53.21	1966	52.92	1976	52.50	1986	54.53	1996	53.22	2006	55.79
1957	53.92	1967	52.66	1977	53.21	1987	54.75	1997	54.99	2007	54.04
1958	51.38	1968	53.23	1978	52.29	1988	53.39	1998	55.18	2008	54.54
1959	54.26	1969	53.98	1979	53.46	1989	53.35	1999	54.34	2009	53.03
1960	52.91	1970	52.97	1980	54.17	1990	56.38	2000	53.15		
1961	53.08	1971	54.07	1981	53.54	1991	56.19	2001	57.55		
1962	51.45	1972	53.34	1982	53.47	1992	53.11	2002	55.94		
1963	51.84	1973	54.87	1983	54.68	1993	54.19	2003	53.37		
1964	52.81	1974	53.71	1984	53.12	1994	54.10	2004	53.62		

For your convenience, the data can also be download as a plain text file from here:

`http://vanderbei.princeton.edu/307/homework/yearlyTemps.txt`

Visually, there appears to be a small upward trend in the average temperatures. Let's try to quantify this trend. To do this, we will try to draw the best straight line through the data:

$$y = m(x - 1955) + b$$

Here, x is the time variable (in years), y denotes the estimated average temperature, m is the slope, i.e., the upward (or downward) trend, and b is the "intercept" (at year 1955). We have $n = 55$ specific data points: (x_i, y_i), $i = 1, 2, \ldots, n$. These data points don't lie exactly on the idealized straight line. Let's denote the deviations by ε_i:

$$\varepsilon_i = y_i - (m(x_i - 1955) + b)$$

Our goal is to minimize the sum of the absolute values of the ε_i's:

$$\text{minimize} \sum_{i=1}^{n} |\varepsilon_i|.$$

As discussed in the context of the Markowitz model, the absolute values can be handled by introducing new variables, $\alpha_i \geq |\varepsilon_i|$, and solving the following linear programming problem:

$$\text{minimize} \sum_{i=1}^{n} \alpha_i$$

subject to the constraints

$$\begin{aligned} \varepsilon_i &= y_i - (m(x_i - 1955) + b) & i &= 1, 2, \ldots, n \\ \varepsilon_i &\leq \alpha_i & i &= 1, 2, \ldots, n \\ -\alpha_i &\leq \varepsilon_i & i &= 1, 2, \ldots, n \end{aligned}$$

To summarize, the data for this problem are the x_i's and y_i's and the variables are the slope m, the intercept b, the deviations ε_i's, and their absolute values α_i's.

1. Write an AMPL model to solve for these variables. Of course, we are mainly interested in the slope m.
   (a) What is the slope in *degrees Fahrenheit per century*?
   (b) What is the estimated average temperature for the year 2019?
2. Change your AMPL model so that it minimizes the sum of the squares of the α_i's.
   (a) What is the slope in *degrees Fahrenheit per century*?
   (b) What is the estimated average temperature for the year 2019?

Note: Submit via Blackboard both AMPL models and the answers to the questions above. With the AMPL models provide a step-by-step explanation of what it is doing and how.