Project 1: Air pollution, background information

These data were collected to evaluate the relationship between air pollution and mortality. The data are from 75 US cities (actually, Census Bureau Metropolitan Statistical Area’s (MSAs), which are a city and the closely linked area surrounding a city). These are a subset of the 382 MSAs currently defined by the US government. You can call the observations cities or MSAs.

For project 1, we will not consider mortality. Instead, you will look at the relationship between an air pollution variable and a single climate variable.

Each row of data represents one MSA. The variables are:

JulyTemp: average July temperature, in degrees Fahrenheit

averaged over the preceeding 30 years (usual Weather Service practice)

HC: annual average daily HydroCarbon concentration in the air

SO2: annual average daily Sulphur Dioxide concentration in the air

Your data set will have either HC or SO2, but not both.

Both HC and SO2 are unitless because of the way they are calculated. Larger numbers mean the air is more polluted.

Specific items to consider are:

1) Describe the relationship between July temp in a MSA and the pollutant concentration.

In other words: A MSA that is 10 degrees warmer has # units more (or X times more) HC, on average.

If looking at HC:

2) You are looking for a quick way to identify new MSA’s that are reasonably likely to exceed 50 units of HC. July temperature values are widely available. What JulyTemp corresponds to a predicted HC above 50? (No need to figure out a precision or confidence interval for the estimated JulyTemp).

If looking at SO2:

2) The six MSA’s with a SO2 value of 1 are the six southern-most MSA’s in the data set (In TX, FL, LA). If these MSA’s are removed from the data set, what can you say about the relationship between July temp and SO2 concentration “without the southern-most cities”?