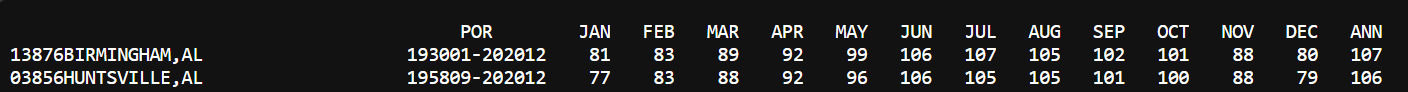
# Homework problem – Learning SAS - Week 2

**Step 1 – Read input data 1:**

Read the following file into a SAS dataset named **work.temps\_high**. This file has the highest temps for the given place for various months.

<https://www.ncei.noaa.gov/pub/data/ccd-data/hghtmp20.dat>

The data is laid out like this.



Read the data as follows.

1. Ignore the header.
2. Read the numeric values in the first 5 columns into a variable called code. Do not discard the leading zeros.
3. Read the name of the city and state into a variable called place.
4. Read the data for columns JAN through DEC into columns named High\_JAN, High\_FEB… through High\_DEC respectively.

**Step 2 – read input data 2:**

Next read the following file into a SAS dataset named **work.temps\_low**. This file has the lowest temps for the given place for various months.

<https://www.ncei.noaa.gov/pub/data/ccd-data/lowtmp20.dat>

the data is laid out like the file above.



Read the data as follows.

1. Ignore the header.
2. Read the numeric values in the first 5 columns into a variable called code. Do not discard the leading zeros.
3. Read the name of the city and state into a variable called place.
4. Read the data for columns JAN through DEC into columns named Low\_JAN, Low\_FEB… through Low\_DEC respectively.

**Step 3 – Combine data and calculate columns:**

Calculate the range for the temps for each place and month.

For example: If Birmingham, AL has a temp of 81 for High\_JAN (in temps\_high dataset) and -6 for Low\_JAN (in temps\_low dataset).

Therefore Range\_JAN should be calculated as 81 - -6 = 87.

Calculate the Range\_XXX for all 12 months (from Range\_JAN …. Range\_DEC).

Range = max -min for each month.

Name the output dataset as **temps\_range**. This dataset should contain only the code, place and 12 range columns (Range\_JAN…. Range\_DEC)

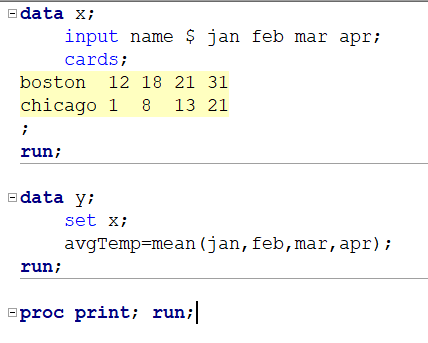
Note: Discard the high\_XXX and low\_XXX variables after calculating the range.

Hint: You might need to sort the 2 datasets (temps\_high and temps\_low) and then merge using the datastep. Then calculate the range variable as high – low for each month.

**Step 4: BONUS**

1. Calculate the average temp range for each city (across all months) and save the value in **avg\_range**, then,
2. Sort the data by ascending average range values and
3. Print the top 10 cities (with lowest fluctuation in temps). Add a title to the print output as ‘Top 10 cities with lowest temp fluctuations through out the year’.

Sample program for calculating mean across variables using the **mean function** in SAS



Output:

