Introduction:

The questions of interest here is to compare and contrast data from two different data sets. The first data set is a series of over 570,000 temperature observations from countries all over the world. The second data set contains over 200,000 observations from 99 large cities around the world. For each data set we analyzed the differences in maximum and minimum temperatures from Janurary 1, 1900 to December 31, 2012. We also created a subset of the temperatures recorded in the US for the previously described time period. We will produce a graph showing the annual change in average temperature over time, as well as show when the greatest year to year change occurred. The final piece of explanatory analysis was to demonstrate graphically the greatest temperature differential for each of the 99 cities in the city data set and graph that difference for the top 20 cities who had the greatest difference. We will then compare the city data chart to the country data chart and comment on the findings.

Code Explanations:

Tim – If you want to do this for your parts, I’ll do it for mine.

City Temp Data:

CityTemperature.R

Here we are importing the raw data from our working directory and putting it into a data frame for ease of manipulation in subsequent steps.

CityDataCleanup.R

This section takes the raw data, changes the date formats to a consistent format and then creates a new data frame with the dates specified in the question of interest. We also change a variable name to make it easier to work with as we begin to summarize and manipulate the data.

CityMaxandMin.R

Short section to calculate Maximum and Minimum temperatures for each city in the data set, and then create an output data frame to generate the graphics from.

Top20Cities\_Plot.R

Generate the graphics indicated in the question of interest section.