**CDS 130 - 001 -- Fall 2011**
**Computing for Scientists**

**Project : Global Temperature Study**

**Assignment Date: Nov. 29, 2010**
**Due Date: Dec. 08, 2010**

## 1. Introduction

The objectives of this project are to (1) learn how to visualize a large scale, two-dimensional scientific data set, and (2) analyze the data using the scientific methods of statistical measure, histogram and linear regression.

## 2. Data and Data Reading

The data provided to you are global long term temperatures gathered by National Centers for Environmental Prediction (NCEP). These data specify the temperature (in degrees Celsius) at every grid point around the globe, for every latitude from South 90 degree to North 90 degree, and every longitude from zero degrees to 360 degrees. Therefore, there are in total 180 X 360 = 64,800 data points in the provided data file.

The data file can be downloaded from the class website at:
http://solar.gmu.edu/teaching/2011_CDS130/project/temp_ascii.txt
You will need to save the file into Matlab's working directory. In the file, the data values are stored in ASCII format in a single column. To read the data into a Matlab program:

```
data=dlmread('temp_ascii.txt',',') ;    %generate 1-D array with 64800 elements
```

To convert the 1-D data array into a 2-D data array, with rows representing latitude and columns representing longitude, do the following

```
for i=[1:180]
  for j=[1:360]
        t(i,j)=data((i-1)*360+j);
  end
end
```

Now, you have the data in 1-D format, stored in the variable "data", and in 2-D format, stored in the variable "t". Check the dimensional size of these two variables using Matlab's "size" command

```
size(t)
```

## 3. Task 1 - Visualization

You need to use visualization tools to visualize the global property of the data:

1. First, use the surface map or 3-D height plot "surf" to visualize the 2-D temperature map.
2. Second, use the Matlab function "image" to plot a colormap of the global 2-D temperature data. Make sure that you choose appropriate color table and color table size to visualize the full range of the data.

## 4. Task 2 - Statistical Measures

Find the following statistical measures of the temperature data: minimum, maximum, mean or average, variance and standard deviation.

## 5. Task 3- Data Analysis

1. First, use the Matlab function "hist" to plot a histogram of the data. The bin data range is -40 to +40 degrees, and each bin size is 5 degrees. From the histogram plot, find the temperature bin of the highest frequency.
2. Second, find the average temperature over all longitudes **for each latitude,** from South 90 degree (matrix "t" row 1) to North 90 degree (matrix "t" row 180). This results in a new 1-D data array of 180 points (say, stored in variable "t_ave"), indicating how temperature changes with latitude. Using this newly created 1-D data array, plot ("plot" in matlab) average temperature versus latitude.
3. Third, make a linear correlation analysis of temperature versus latitude. The latitude can be obtained by converting the row index I as follows (the latitude variable's range is -90 to 90):

```
lat = [1:180]-90
```

In order to make a correlation analysis between the temperatures and the latitudes, you need to convert the signed latitudes into absolute latitudes. This can be done as follows:

```
lat_abs = abs(lat)
```

Now, you have two new variables: "lat_abs" and "t_ave".

You are required to use the Matlabfunction "polyfit" to find the **linear** correlation function between the data stored in "lat_abs" and "t_ave" (linear means that n = 1 when using "polyfit"). Furthermore, you're also required to use the Matlab function "corrcoef" to find the correlation coefficient.

Finally, you're required to make a linear correlation plot, showing the scattering of the data points (average temperature vs. latitude), and the fitted line.

## 6. Submission

You are required to write a succinct project report summarizing your results, as required by the tasks 1, 2 and 3. The report should include all the images that need to be plotted. Your Matlab program code should be attached to the report. I only accept the electronic copy of the project report. **Send me the report in a single attached file (either in DOC format or PDF format) by e-mail (jzhang7@gmu.edu)**

## 7. Project Grading:

The total grade of the project is 100 points: (1) (30 pts) Overall writing of the project, including the clarity of the writing and the sufficient discussion of the results (2) (15 pts) the success of task 1, (3) (15 pts) the success of task 2, (4) (20 pts) the success of task 3, (5) (10 pts) proper labeling of the figures. Each figure should have xlable, ylable, title and legend, and (6) (10 pts) readability and proper documentation of the code.

**Bonus task 1 (5 pts):**

What is the temperature in Fairfax, Virginia, as inferred from the data?

**Bonus task 2 (5 pts):**

Where is the location of the highest temperature, in terms of longitude and latitude?