EEN118 LAB NINE

This lab is all about data processing. You will be reading information from a file that contains thousands of individual pieces of data. There is far too much for a person to deal with in its numeric form, but you will display it in a way that is very easy to understand.

On the class web site, associated with this lab, there are 76 data files, each containing a whole year of weather observations from a different location in the U.S. Choose one of them and down-load it to the computer you are using. Sadly all the data is from 2003, the last year it was made public in a form like this.

The files all have one line of data for each day of a particular year, and each line consists of exactly nine numbers. But beware: this is real data from real meteorological stations. Sometimes their equipment isn't in perfect working order, so some days might not appear.

As an example of the data, here are three lines taken from near the end of the file for Mount Washington, NH.

```
2003 12 16 -17.7 -7.7 0.0 99 3 34.4
2003 12 17 -3.8 -1.1 1.1 99 -1 24.5
2003 12 18 -13.3 -8.3 0.0 130 53 36.6
```

The first three numbers on each line give the year, month, and date of the observations. The remaining numbers are:

4. The Minimum Temperature recorded on that day (Fahrenheit)
5. The Average Temperature recorded during that day (Fahrenheit)
6. The Maximum Temperature recorded on that day (Fahrenheit)
7. The Depth of new Snow fall (tenths of an inch)
8. Total amount of Precipitation, incl. snow melted (tenths of an inch)
9. The Maximum Wind speed observed (miles per hour)

If you were wondering, their rain meter was broken on 17th December. That's what `-1` indicates in the last three items.

1. *Make sure you can read the file*

Write a simple program that reads the whole file, and gives you just enough information to verify that it is working properly. Perhaps you could just print out the average temperature for every day, and check that it does generally get warmer in the first half of the year and colder towards as the end approaches.

As a reminder, here is a little snippet of code that would open a file, read three numbers from it, then close it again.

```
ifstream fin("file.name");
int a, b, c;
fin >> a >> b >> c;
if (fin.fail())
   cout << "Not enough data in the file!\n";
else
   cout << "the average of the numbers was" << (a + b + c) / 3 << "\n";
fin.close();
```

If your program can't find the file:

If you get totally wrong information, it may be that you didn't download the file to the right place, or maybe got the file name wrong. First, find out if that sort of thing really is the reason. Add a bit more code directly after opening the file:

```
ifstream fin("file.name");
if (fin.fail())
{ cout << "The file didn't open\n";
  exit(); }
```

If you see that the file didn't open, it is very likely to be because it is in the wrong place. To find the right place, create a file with an easy to spot name:

```
ofstream fout("very-easy-name-to-notice-XXXXXXXXXX.txt");
fout << "X\n";
fout.close();
```

After running that, just search for the file with that name, it will be in your project folders somewhere. The input file that you want to read should be put in exactly the same folder.

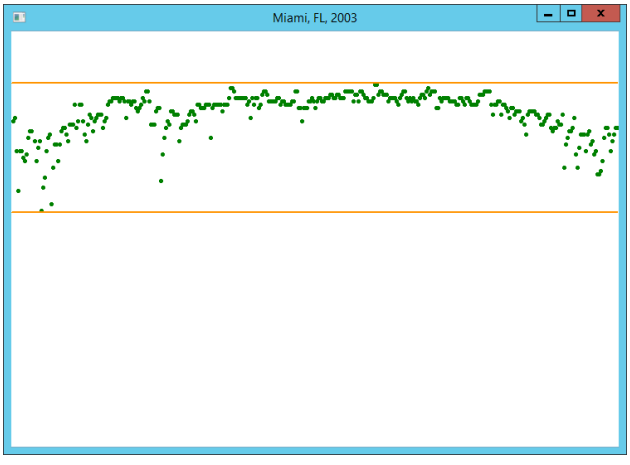
Another way is to download the file to anywhere you like, but carefully note down the exact location you chose, then use that as the beginning of the file name:

```
ifstream fin("C:\\downloads\\MIAMI-FL.txt");
```

That should be fool-proof if you remember two things: the \'s used to separate folder names under windows need to be doubled, and under some systems it matters whether letters in a file name are capital or little. That last one shouldn't apply to windows systems, but better safe than sorry.

2. *Turn the numbers into a graph*

Adapt your program so that it displays all the average temperatures as points on a graph. Pick a reasonably large window and pen size so that it will look OK. The vertical position is of course just the average temperature multiplied by some suitable scaling factor, but what about the horizontal position?



Make it easy on yourself, just for now. If you pretend that every month is 31 days long, then `month * 31 + day` gives a simple basis for the x-coordinate, which you can scale or shift as desired.

Add two horizontal lines, showing the minimum and maximum average temperature recorded for the whole year.

The picture above was produced from the Miami FL file. If you want to set the caption of the window like I did, there is a handy little function called `set_caption`. It takes a string parameter.

3. *Get the X coordinate right.*

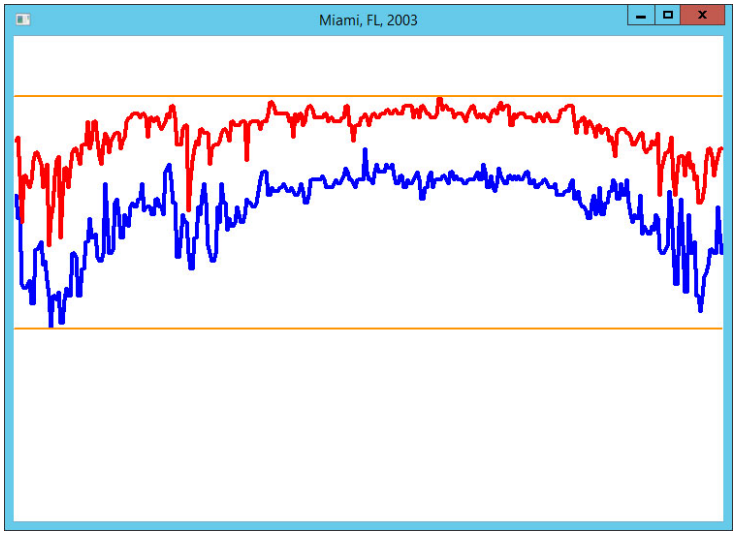
There aren't really 31 days in every month, and pretending that there are can cause annoying gaps. Fortunately, you recently had an assignment that involved working out what day of the year a particular date is. Make it so that the horizontal position of each dot is correct: the distance between 28th February and 1st March is the same as the distance between 1st March and 2nd March, or any other two consecutive days.

4. *Less dottiness*

In the summer, the temperature tends not to change so much from day to day, so the dots almost merge into a solid line. In winter, as you can see, it is harder to see what is going on. There are dots all over the place. Make it into a line graph instead.

5. *More information*

When someone is planning a trip and deciding what to pack, knowing the minimum and maximum temperature is much more useful than just the average. Improve your program so that it shows a red line graph for the daily maximum temperature, and a blue line graph for the daily minimum. On the same axes of course.

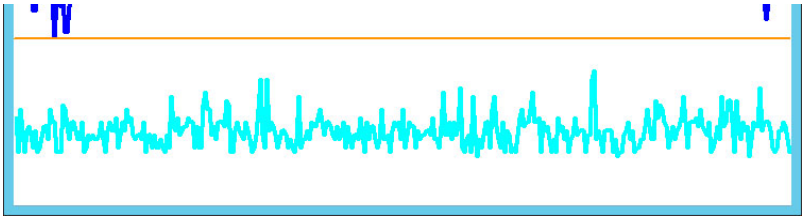


This is one of those situations where the restrained use of a few variables can be really useful. Be sure that you only read through the file once, even to plot two lines.

6. *Kites*

We also want to know whether there will be a nice cooling breeze in the summer, or an unpleasant wind in the winter, or just enough for flying a kite. For most places, if you plot the day's maximum wind speed at the same scale as the temperatures, the line will stay out of the way at the bottom of the graph.

Add the wind speed to your graph, staying with the requirement that you only read the file once. Sadly, Miami's maximum daily wind speed is fairly consistent, so it doesn't look very interesting. We didn't have a hurricane in 2003.



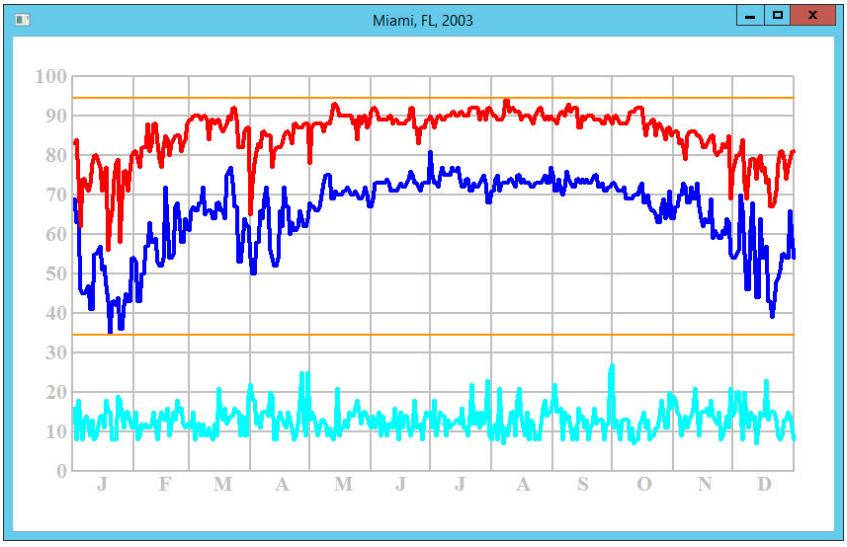
Other places produce more interesting graphs. Look at Mount Washington NH for example.

7. *Make it useful*

It's a nice looking graph, but not really very useful. It's hard to tell just where the last week of October is, and nobody has any idea what the actual temperatures are: just how low was that dip in January?

Add some clearly labelled axes: vertical lines showing where each month starts, and horizontal lines for each ten degrees or so. If you draw them in gray, it won't seem too obtrusive.

This is where good clean programming pays off. If you were keeping things reasonably tidy, it will be easy to work out where the lines go, and to shift the graph to make room for the labels.



8. *Last thing before you go*

Make sure you have written a good robust program. Download a file for a different city, and check that your program displays its data correctly. It is important to make sure that your program works correctly even for data files that have a lot of missing data. One of the data files is for a made-up town called Hopeless, Missouri (at least, I don't think there's a real town with that name, but I didn't check). It was deliberately made with a lot of missing data to help you with testing.

9. **Extra Credit**

There is more information than just temperatures in those files. Allow the user to select which pieces of information they want to see graphs for, and what colours those graphs should be in. You may have to think about the scaling a little.

Adaptive scaling would be nice too. All my samples have assumed that temperatures will fit nicely into a 0 to 100 scale. That is untrue for a great many places. Your program could scan the file first to find out what the true range is, then re-read the data and plot it at a custom-made scale.

Another useful option would be to plot a moving average, which would smooth the graphs out a bit, hiding short fluctuations but showing overall trends more clearly.