AUTHOR: CHARMAINE BONIFACIO
UPDATED: SEPTEMBER 30, 2014
DEPARTMENT OF GEOGRAPHY, UNIVERSITY OF LETHBRIDGE

THE CANADIAN CLIMATE DATA SCRAPING TOOL GUIDE

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Christopher Hugenholtz, Dr. Stefan Kienzle and Thomas Barchyn, who have made their contributions to the Canadian Climate Scraping Tool and user guide documentation.

DISCLAIMER

The Canadian Climate Scraping Tool (CCDST) is an independent project and is not affiliated with, nor has it been authorized, sponsored, or otherwise approved by Microsoft Corporation. The purpose of this document is to provide basic guideline for users interested in CCDST. The macro is provided "as is" without warranty of any kind, either express or implied, including, without limitation, any and all implied suitability for any given purpose or task. Although every effort and care has been taken to make the information as accurate as possible, the author shall not be liable for any error, harm or damage arising from using the instructions given in this documentation or the use of the program.

TRADEMARKS

Microsoft, Visual Basic, Excel and Windows are registered trademarks of Microsoft Corporation in the United States and/or other countries.

The Canadian Climate Data Scraping Tool (CCDST) was developed to expedite the process of obtaining meteorological data from Environment Canada website. Users often expend considerable time not only to manually download but also preprocess long records of historical weather data. Compiling data also requires a considerable amount of time and specific software packages to open specific file type. This tool will allow users to download and compile hourly, daily or monthly data from stations listed on the Environment Canada Climate's National Climate Data and Information Archive. The process of downloading, preprocessing and compiling data is combined together with using 32-bit or 64-bit Microsoft Excel® software under a Windows® operating system.

I. Summary of Tool Operation

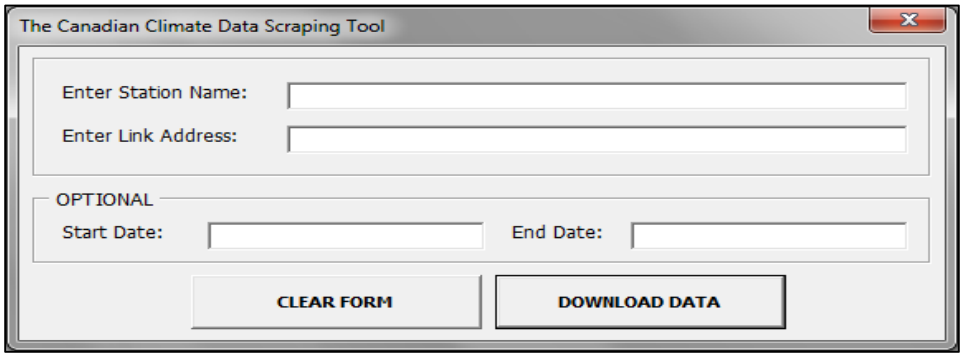

Figure 1. CCDST User Interface

This summary briefly outlines how the tool operates in term of the necessary user input and what the tool will output upon successful completion of the user request.

A. User Input:

- Station Name – user preferred name
- Link address – copied link from Environment Canada using Customized Search (see Section II).
- Start and End Date (optional) – user preferred date range
- Clear form – to reset user input

B. Tool Output:

- Original raw data from Environment Canada
- Preprocessing data from raw data
- A compiled file of the station data

II. How to Operate the Canadian Climate Data Scraping Tool.

1. Ensure a running Windows® operating system with Microsoft Excel® 2007 or newer version and download the macro (that is, if file has not been downloaded).
2. Open the .XLSM file containing the macro that will run the Canadian Climate Data Scraping Tool (CCDST).
3. Once Microsoft Excel is open, ensure that editing, content and macros are enabled.

Figure 2. Security Warning for macro-enabled files.

4. Click on the image found in the worksheet to run the Canadian Climate Data Scraping Tool. The user will be prompted with the following interface.

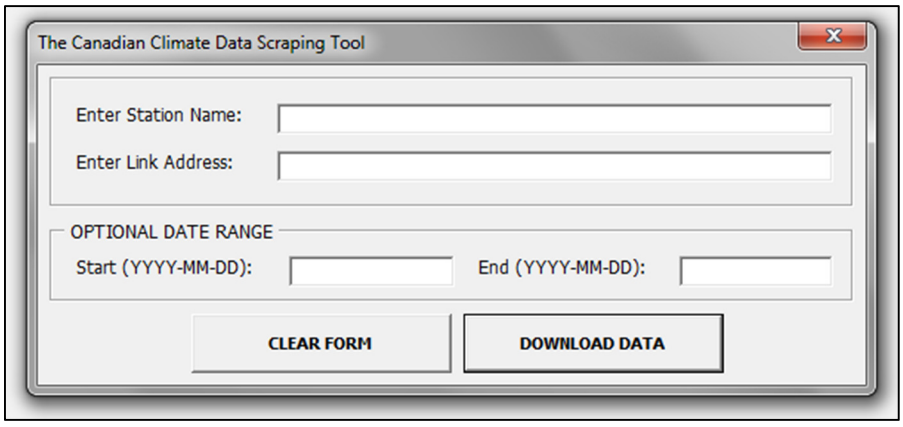

Figure 3. The interface for the Canadian Climate Data Scraping Tool.

5. Click on **Download Data** button. Ensure the following fields are entered (**mandatory**):
    - **STATION NAME:** *Enter the preferred station name. (e.g. Lethbridge A or LethbridgeA are valid station names).*

- **LINK ADDRESS:** *Paste the copied link address from the Environment Canada Climate Data Online website (see Section III on how to obtain the correct link address).*

NOTE: The Canadian Climate Data Scraping Tool (CCDST) only utilizes the ***Customized Search*** *option in the* ***Environment Canada Climate Data Online*** *website. See Section III for more information.*

The following fields are **optional**:

- **DATE RANGE:** *The user has the option to enter the date range in YYYY-MM-DD format. Otherwise, the default date range will be used. The date range must not be out of the scope for what is available on the website.*

6. Select a folder. If no folder is selected, the default file directory location is C:\ drive.

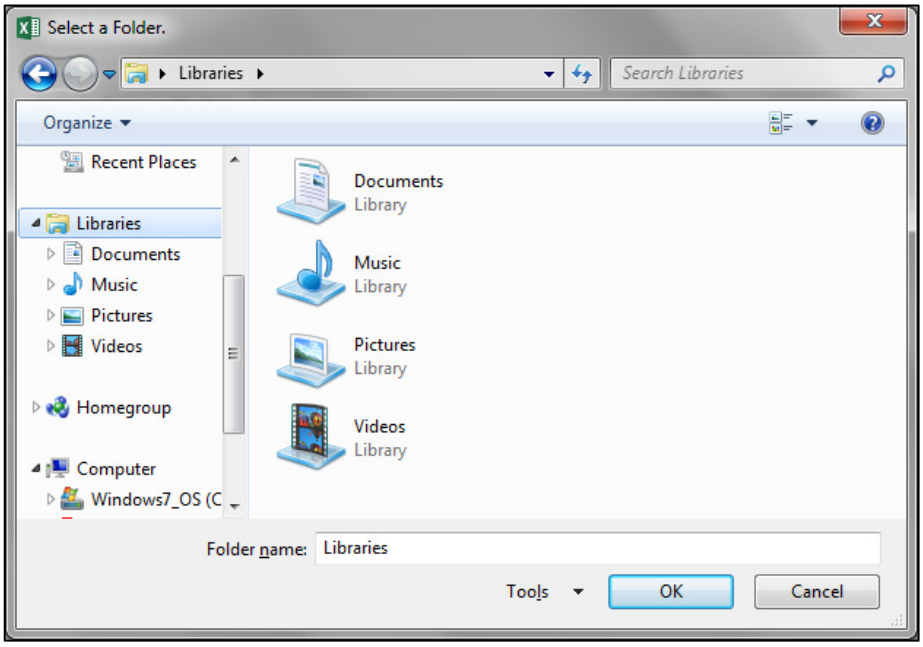

Figure 4. Option to select a folder

7. A Warning Message will appear. Simply click OK or wait until the window closes. Microsoft Excel will be unavailable for other use while processing the user request.

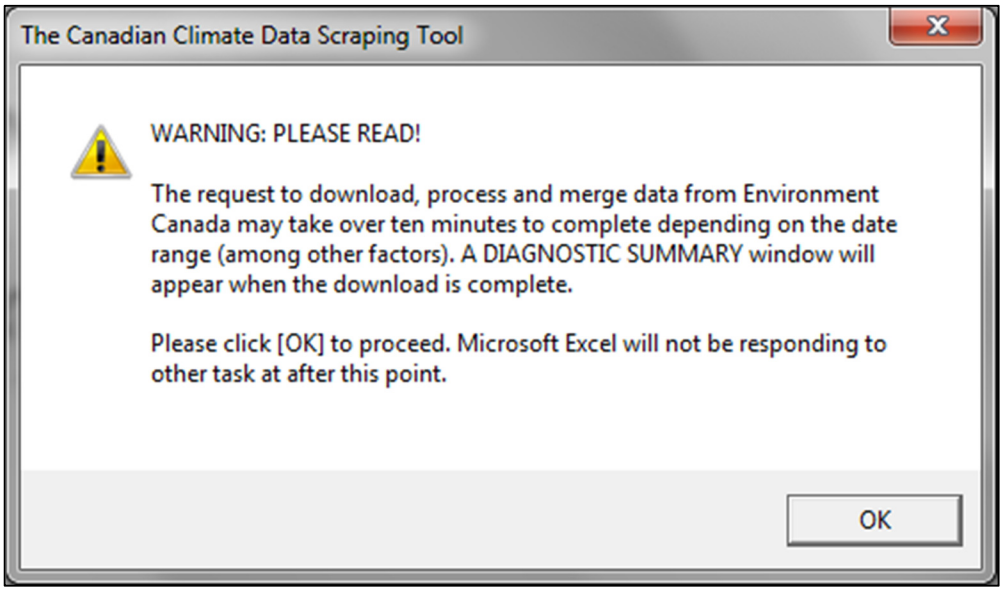

Figure 5. Warning Message will appear

8. Make sure to check the status bar which is located at the bottom left corner.

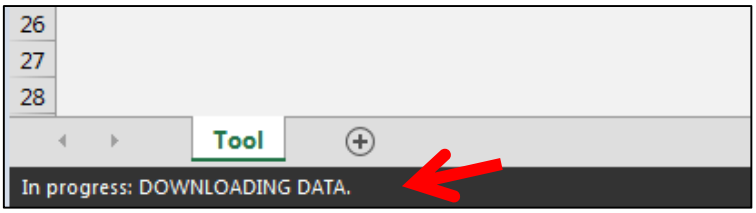

Figure 6. Status Bar Message

9. Upon successful completion of the request, the Diagnostic Summary will prompt the user.

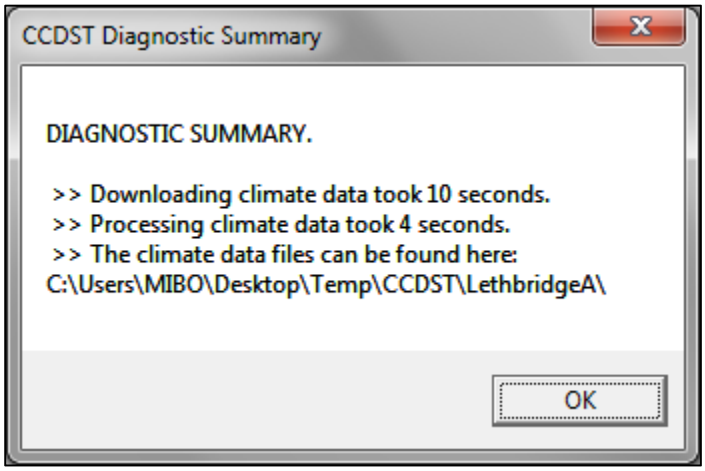

Figure 7. Diagnostic Summary window will appear.

10. Close the tool before proceeding to open any Microsoft Excel files.

The original .CSV files will be saved under
…\CCDST\<StationName>_DOWNLOAD.

The processed .CSV files will be saved under
…\CCDST\<StationName>_PROCESSED.

The compiled .CSV file will be saved under
…\CCDST\<StationName>

III. How to obtain historical weather data using the Customized Search from Environment Canada website.

NOTE: The CCDST was revised due to the recent changes in the Environment Canada Climate Data website.

1. Visit Environment Canada Climate Data Online:

   http://climate.weather.gc.ca

2. Scroll down the page. Enter the station name and click on Search.

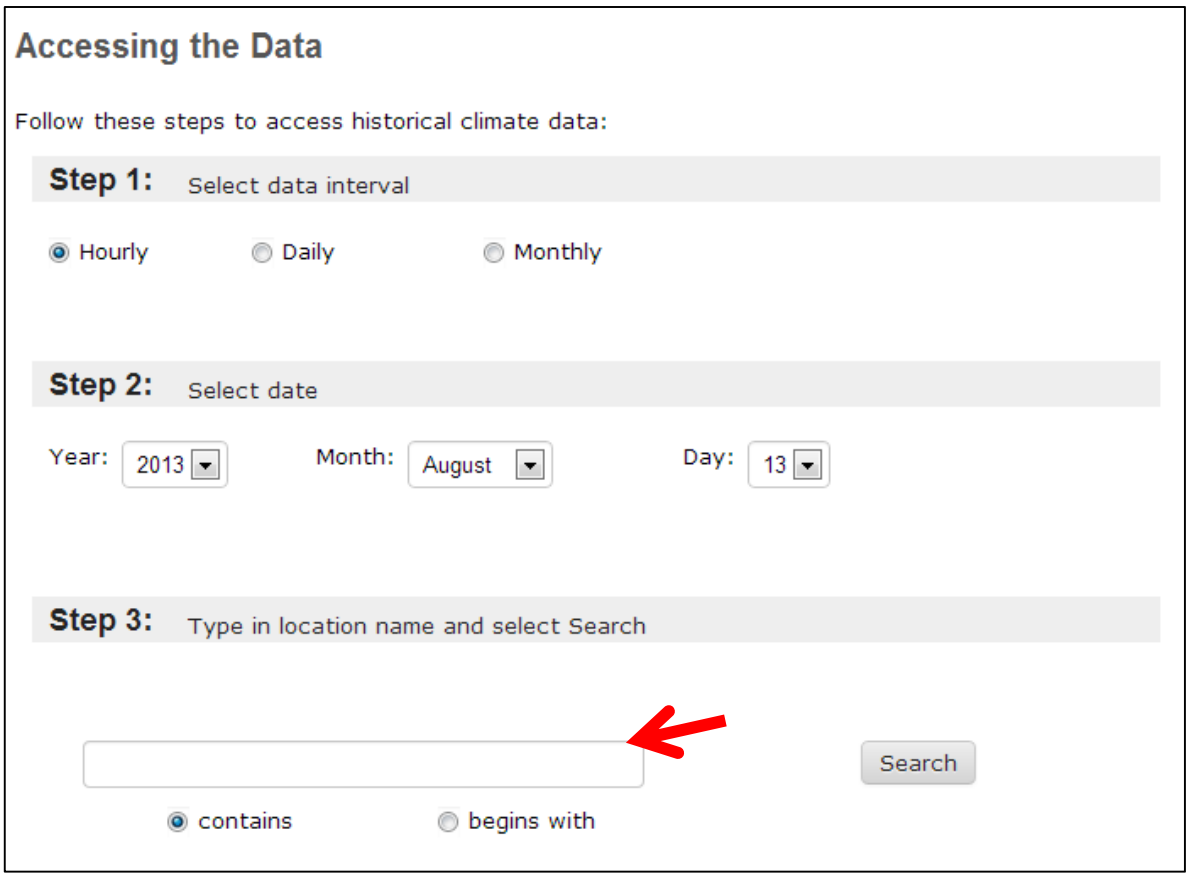

Figure 8. The Environment Canada Climate Data Online website.

3. Or click on Data and select the Advanced Search. Simply Scroll down to the second option, Search by Station Name and enter the name of location you are searching for (e.g. Lethbridge) or keyword (e.g. Leth) and click on Search. This is the preferred method of obtaining link address for each station.

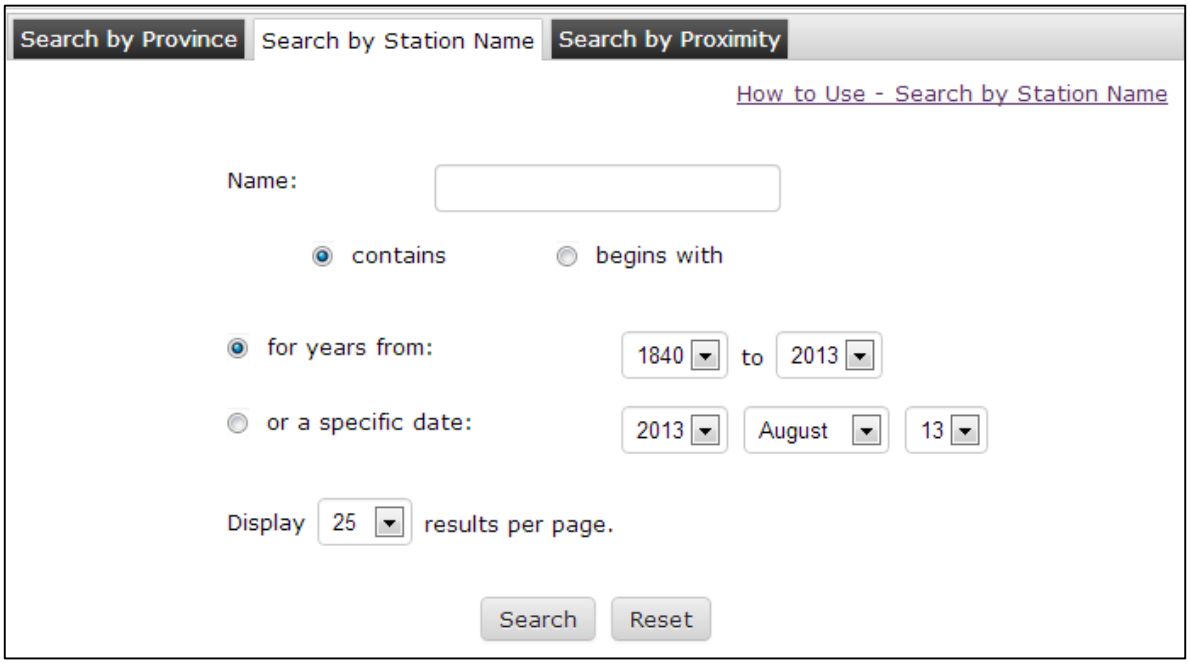

Figure 9. Customized Search Option

4. The Station Results page will be show the list of all the stations that have all of the criteria specified in your search. Select the appropriate station location, data interval, day, month and year, if available. Click on Go.

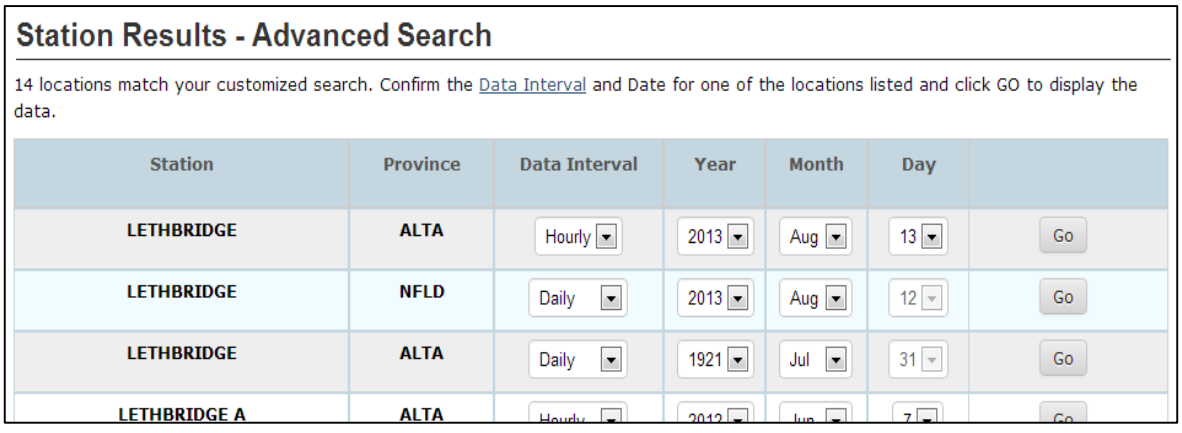

Figure 10. Station Results page.

5. The appropriate data report will be shown according to the criteria specified (e.g. hourly data for Lethbridge, Alberta for August 13, 2013).

Hourly Data Report for August 13, 2013

All times are specified in Local Standard Time (LST). Add 1 hour to adjust for Daylight Saving Time where and when it is observed.

Metadata including Station Name, Province, Latitude, Longitude, Elevation, Climate ID, WMO ID, TC ID

LETHBRIDGE ALBERTA					
Latitude:	49°37'49.000" N	Longitude:	112°47'56.000" W	Elevation:	929.00 m
Climate ID:	3033875	WMO ID:	71267	TC ID:	YQL

Related Data	Additional Search Options	Download Data
Almanac Averages & Extremes (August 13)	Nearby Stations with Data	Hourly Data (August 2013)
Daily Data (August 2013)	Advanced Search	CSV XML
		Download Data

Previous Day | 2013 | August | 13 | Go

Hourly Data Report for August 13, 2013

	Temp °C	Dew Point Temp °C	Rel Hum %	Wind Dir 10's deg	Wind Spd km/h	Visibility km	Stn Press kPa	Hmdx	Wind Chill	Weather
TIME										
00:00	14.6	13.8	95		0	16.1	91.28			NA
01:00	12.8	12.2	96	25	17	8.1	91.31			NA
02:00	13.3	12.8	97	23	9	6.4	91.25			NA
03:00	12.7	12.2	97	31	8	8.1	91.20			NA
04:00	12.1	11.7	98	17	5	6.4	91.21			NA
05:00	11.7	11.6	99	10	13	2.0	91.29			NA
06:00	12.6	12.3	98	18	11	4.8	91.26			NA
07:00	15.0	14.2	95		0	11.3	91.26			NA
08:00	16.8	15.1	90	16	11	16.1	91.27			NA
09:00	18.5	15.1	80	13	9	16.1	91.27			NA

Figure 11. Station Report page.

6. Go to the address bar on any internet browser. Copy the URL address. The URL address is necessary to download all the data in the time series, if available.

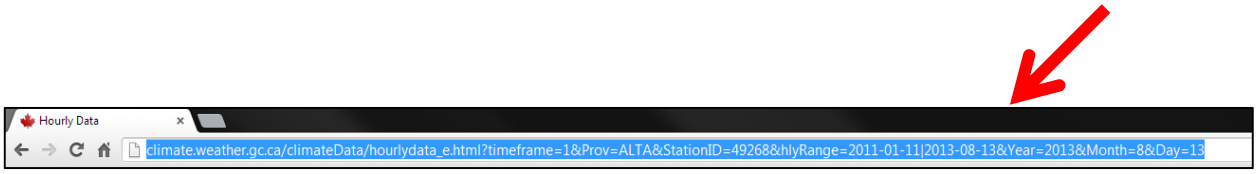

Figure 12. Address Bar

7. Save the URL for another time. Or, proceed to use the Canadian Climate Data Scraping Tool.

NOTE: Link address should look similar to the link address shown below.

- *http://climate.weather.gc.ca/climateData/hourlydata_e.html?timeframe=1&Prov=ALTA&StationID=49268&hlyRange=2011-01-11|2013-08-13&Year=2013&Month=8&Day=13*

- *http://climate.weather.gc.ca/climateData/hourlydata_e.html?timeframe=1&Prov=ALTA&StationID=49268&hlyRange=2011-01-11%7C2014-03-28&Year=2014&Month=1&Day=1&cmdB1=Go#*)

IV. Summary of File Outputs.

A. Raw File Format.

	A	B	C	D	E	F	G
1	Station Na	NAKISKA RIDGETOP					
2	Province	ALBERTA					
3	Latitude	50.94					
4	Longitude	-115.19					
5	Elevation	2543					
6	Climate Id	305MGFF					
7	WMO Ider	71245					
8	TC Identif	WNR					
9							
10	Legend						
11	[Empty]	No Data Available					
12	M	Missing					
13	E	Estimated					
14	B	More Than One Occurrence and Estimated					
15	I	The value displayed is based on incomplete data					
16	S	More than One Occurrence					
17	T	Trace					

Figure 13. This is a sample raw file. The file will contain headers that may not be necessary for data analysis.

B. Processed and Compiled File Format.

	A	B	C	D	E	F	G	H	I	J	K
1	YEAR	MONTH	MEAN_MAXTMP_C	MEAN_MAXTMP_F	MEANTMP_C	MEANTMP_F	MEAN_MINTMP_C	MEAN_MINTMP_F	EXTR_MAXTMP_C	EXTR_MAXTMP_F	EXTR_MINTMP_C
2	1999	4	-1.6	E					7.7		
3	1999	5	1.4						14.2		
4	1999	6	6.7	E	3.5	E	0.3	E	15.4	E	-5
5	1999	7	9.4		5.9		2.4		17.7		-7.2
6	1999	8	12.9		9.5		6		19.9		-3.3
7	1999	9	7.3	E	3.7	E	0.1	E	17.1		-8.2
8	1999	10	1.2		-1.7		-4.5		10.8		-14.6
9	1999	11	-1.5	E	-4.2	E	-6.9	E	8.6		-15.2
10	1999	12	-2.8		-5.5		-8.2		9.3		-20.6
11	2000	1	-7.6		-10.2		-12.8		2.5		-20.1
12	2000	2	-5.4	E	-8.4	E	-11.4	E	5.4		-20.8
13	2000	3	-4.7	E	-7.4	E	-10.1	E	1.7		-16.3
14	2000	4	-0.2		-4		-7.8		8		-23.4
15	2000	5	2.7		-0.4		-3.4		8.8		-7.9
16	2000	6									

Figure 14. This is a sample processed file. The file will only contain the column headers and the data from Environment Canada. The file is the final product that can be used in other software packages (e.g. R or Matlab).

End of Guide