Downloading Data from Environment Canada

Website:

<http://climate.weather.gc.ca/historical_data/search_historic_data_e.html>

Step 1: Find Station IDs

Download StationInventoryEN.csv from:

<ftp://ftp.tor.ec.gc.ca/Pub/Get_More_Data_Plus_de_donnees/>

This file has a list of all weather stations in Canada, their station IDs, and their years of operation. From this list, decide what stations you are interested in and collect their IDs. See Readme.txt at this website for more information.

Step 2: Make List of URLs

Create a .txt file that contains a list of URLs according to the following form:

http://climate.weather.gc.ca/climate\_data/bulk\_data\_e.html?format=csv&stationID=2383&Year=${year}&Month=${month}&Day=${day}&timeframe=1&submit=Download+Data

Each line in the text file should be a different URL. This URL has several main parts that one can edit:

i) http://climate.weather.gc.ca/climate\_data/bulk\_data\_e.html?format=csv

Leave this part as it is.

ii) &stationID=**2383**

The number in bold is the station ID. Change the number to match the ID of whatever station you’re looking at.

iii) &Year=${year}

Change ${year} to be the year of interest.

iv) &Month=${month}

Change ${month} to be the month of data that you want to look at, if looking at month-long series of data (as opposed to year-long data). If you are looking at all months in the year, change ${month} to be any number (ie: &month=1).

v) &Day=${day}

Change ${day} to be the day of data that you want to look at, if looking at a single day-long series of data. If you are looking at month- or year-long series of data, change ${day} to be any number (ie: &day=1).

vi) &timeframe=1

This describes the type of data you are looking at.

&timeframe=1 (hourly data)

&timeframe=2 (daily data)

&timeframe=3 (monthly data)

vii) &submit=Download+Data

Don’t touch this. It tells the computer to actually download the file.

URL Example:

Description of data:

Station: 26

Year: 2010 (whole year)

Month: all

Day: all

Type: daily

Corresponding URL:

http://climate.weather.gc.ca/climate\_data/bulk\_data\_e.html?format=csv&stationID=26&Year=2016&Month=1&Day=14&timeframe=2&submit=Download+Data

URL List Example

Description of data:

Stations: 26 and 2383

Years: 2000-2002

Month: all

Day: all

Type: Daily

Corresponding list of URLs:

http://climate.weather.gc.ca/climate\_data/bulk\_data\_e.html?format=csv&stationID=26&Year=2000&Month=1&Day=1&timeframe=2&submit=Download+Data

http://climate.weather.gc.ca/climate\_data/bulk\_data\_e.html?format=csv&stationID=26&Year=2001&Month=1&Day=1&timeframe=2&submit=Download+Data

http://climate.weather.gc.ca/climate\_data/bulk\_data\_e.html?format=csv&stationID=26&Year=2002&Month=1&Day=1&timeframe=2&submit=Download+Data

http://climate.weather.gc.ca/climate\_data/bulk\_data\_e.html?format=csv&stationID=2383&Year=2000&Month=1&Day=1&timeframe=2&submit=Download+Data

http://climate.weather.gc.ca/climate\_data/bulk\_data\_e.html?format=csv&stationID=2383&Year=2001&Month=1&Day=1&timeframe=2&submit=Download+Data

http://climate.weather.gc.ca/climate\_data/bulk\_data\_e.html?format=csv&stationID=2383&Year=2002&Month=1&Day=1&timeframe=2&submit=Download+Data

Step 3: Download data in bash

Open terminal and go to the directory that you want to save your files in. Then, in the command line, type:

wget –-content-disposition –i URLs.txt

Here, URLs.txt is your list of URLs. This will download the file specified by each URL to its own file in the current directory. If you want to save all URLs to the same specified file, enter the following in the command line:

wget –-content-disposition –i URLs.txt –O output.csv

Here, output.csv is the name of the output file, and –O is the capital letter ‘Oh’ as in ‘Omega’, not a zero. Note that this file will then need to be processed, as each data set will have its own header.

If you want to download data from a single URL (say, the data from URL example above), use the command

wget --content-disposition "http://climate.weather.gc.ca/climate\_data/bulk\_data\_e.html?format=csv&stationID=26&Year=2010&Month=1&Day=1&timeframe=2&submit=Download+Data"

You can also add –O output.csv if you want to specify the output filename.

Note:

Don’t copy and paste URLs from here to test them out (Word’s formatting gave me issues when I tried this, so just write out the URLs that you want directly).