Use of FILENAME statement to create automated reports with severe weather events data from the National Weather Service

Valentin Todorov, Assurant Specialty Property, Atlanta, GA

Abstract

Risk managers at insurance companies and banks constantly monitor financial exposure and potential impact from various weather related events – hurricanes, tornadoes, wind and hail. To understand the exposure to such phenomena, managers can utilize the severe weather events data from NOAA's National Weather Service Center (http://www.spc.noaa.gov/climo/reports/). This paper shows how the FILENAME statement can be employed to easily obtain and analyze this information. The Weather Service publishes daily files containing information about severe weather events on the NOAA's website. The files are not cumulative, which makes the aggregation of the information very tedious. With the help of the FILENAME statement it is easy to obtain and aggregate files for as many periods as needed. The aggregated data can then be used for various purposes: 1) Find properties in companies' portfolio, close to a tornado touchdown to estimate financial impact; 2) Determine areas with high likelihood to incur financial losses from severe events; 3) Create custom reports and maps. The material from this presentation can be extended to obtain content from other sources (e.g., Census Bureau, Bureau of Labor Statistics).

Introduction

Damages from weather related events (hurricanes, tornadoes, wind and hail) cost billions of dollars annually to insurance companies and banks. To understand the impact from such weather phenomena, loan portfolio and risk managers pay close attention to severe weather events, their location of occurrence and strength.

The National Weather Service, a unit of the National Oceanic and Atmospheric Administration (NOAA), provides daily updates of severe weather events on its website. The information can be downloaded for free and used for analysis. Since NOAA provides one file per day it is very time consuming to collect and summarize the data for a few months at a time. However, with the help of the FILENAME statement the task is easy to accomplish[1].

FILENAME Statement in a Nutshell

The FILENAME statement is used to associate a fileref with a file or a device. Once assigned, the referenced information can be used in SAS ® or imported, if the information is not in a format native to SAS.

The syntax of the FILENAME is:

FILENAME fileref <device-type> 'external-file' <ENCODING='encoding-value'> <options><operating-environment-options>;

---

[1] When using the FILENAME statement to access information outside of your corporate network, you may need to ask the system administrator to allow SAS to access external sources.

To access information, the statement provides a wide range of methods, which are broadly categorized into two groups. In the first group we can put methods used to access information located on a physical device which can be a local hard disk, remotely on a server or a web site. In the second group are methods that can be used to access information that doesn't reside on a physical device, e.g. catalog and clipboard. Interested readers should refer to DeGuire (2007) for a detailed description of the FILENAME statement.

Severe Weather Events Database at the National Weather Service
The severe weather events database at the National Weather Service collects historical information about weather events that occur throughout USA. The data is released daily in comma delimited files (CSV) and can be downloaded from the NOAA's website. There are three severe weather events for which data is provided – Tornado, Wind and Hail. The file names are in the format *[date event occurred]_rpts_[type of event].csv*

For example, the link to the file with tornado occurrences on April 30, 2010 is:
http://www.spc.noaa.gov/climo/reports/100430_rpts_torn.csv

Similarly, the link to the file with hail data for April 30, 2010 is:
http://www.spc.noaa.gov/climo/reports/100430_rpts_hail.csv

The files provided by NOAA contain information about the time when the event occurred, geographical location, as well GIS coordinates (latitude and longitude of the location), which can be used to map the events. Additionally, there is a field with narrative information about the strength of the tornado, injured people and anything else that might be of interest to the public.

NOAA Macro Implementation
The purpose of the macro is to import the daily report files from NOAA, clean and concatenate each one and combine them into a single file that can be used for analysis. Below I have described in detail how the macro achieves this.

Step 1. Choose the type of severe weather event for which data is needed and define the start and end dates of the reporting period. The dates entered can have various formats, because the macro converts the entered dates to SAS dates using the ANYDTDTE9. format at Step 4 below.

```
** Provide initial parameters for the report;
%let event  = torn; /*Weather events – torn (for tornado), wind, hail*/
%let start_date   = 01Aug2009;
%let end_date     = 30Jun2010;
%let library      = work;
```

Step 2. Create an empty file that will serve as a placeholder to which all files imported by the macro will be appended later.

```
** Delete the dataset from a previous report;
proc datasets lib=&library.;
delete &event._append;
run;

** Create new placeholder to which data will be appended;
data &library..&event._append;
input
_F_Scale      $3.
Speed         $3.
Size          8.
_Location     $20.
_County       $13.
_Comments     $230.
date          8.
Time          8.
State         $2.
Latitude      8.
Longitude     8.
event         $10.
;
format date mmddyy10.;
datalines;
;
run;
```

Step 3. Define the macro and name it NOAA. It is preferred to add the system option MPRINT, which displays the statements generated by the macro execution. The information can be used to debug the macro, if problems occur with its execution.

```
** Define the macro call named NOAA;
options mprint;
%macro noaa(start,end);
```

Step 4. To allow for flexibility of the dates entered at Step 1, the macro reformats the start and end dates using the ANYDTDTE9. date format. After reformatting the dates, the macro calculates the number of days between the start and end dates and creates a DO LOOP with the number of iterations equal to the difference between those dates. In the first iteration of the loop the macro uses a report date equal to the start date, but in each consecutive iteration the date is advanced by one day forward. In the last iteration the macro uses report date equal to the end date. The advancement of dates is accomplished with the INTNX function.

```
** Convert the dates to SAS date;
%let start=%sysfunc(inputn(&start,anydtdte9.));
%let end=%sysfunc(inputn(&end,anydtdte9.));

** Determine the number of days between the start and end dates;
%let dif_dates=%eval(&end-&start);

%do i=0 %to &dif_dates;

** Advance the date by one day starting with the start date;
%let date=%sysfunc(putn(%sysfunc(intnx(day,&start,&i,b)),yymmdd6.));
```

Step 5. The macro starts off with referencing the file located on the NOAA's website, which is accomplished with the FILENAME statement. The method used to access the file is HTTP. The reference to the file is called NOAAfile, which is then passed to the PROC IMPORT. After the file is imported, the data fields are formatted with the same length, because at the last step all files are concatenated in the final output dataset. On days with no weather events, no files will be imported and SAS will output spurious messages to the log. To prevent those messages, the macro checks for the existence of files using the EXIST function. If a file exists, the data step is executed and a file is appended to the final dataset, otherwise a warning is written to the log.

```
** Reference the report file;
filename NOAAfile HTTP
"http://www.spc.noaa.gov:80/climo/reports/&date._rpts_&event..csv";

** Import the file;
proc import datafile    = "NOAAfile"
            out         = &library..&event._&date.
            dbms        = CSV
            replace;
run;

** Check for existence of the imported file;
%if %sysfunc(exist(&library..&event._&date.)) %then %do;

    ** Modify the imported file;
    data &library..&event._&date.;
    set &library..&event._&date.;
    length _F_Scale $3. _Location $20. _County $13. _Comments $230.
    event $10.;
    _F_Scale    = F_Scale;
    _Location   = Location;
    _County         = County;
    _Comments   = Comments;
    Latitude    = Lat;
    Longitude   = Lon;
    drop F_Scale Location County Comments Lon Lat;

    ** Assign the date when an event has occured;
    date = input("&date",yymmdd6.);
    format date mmddyy10.;

    ** Assign the event type;
    event = propcase("&event");
    if event = 'Torn' then event = 'Tornado';
    run;
```

Step 6. After importing a daily report file, the macro appends it to the empty placeholder file that was created at Step 2. The aggregation is done using the PROC APPEND. It is important to use the option FORCE, because the files contain different fields, depending on the type of severe weather event.

At step 5 the macro used a function to check for the existence of an imported file. If it exists, the file is appended to the final dataset, but if it doesn't a warning message is output to the log.

```
        proc append base=&library..&event._append
              data=&library..&event._&date. force;
        run;
%end;

** If the there is no weather event on a selected date, SAS will not
import a file, but will output a warning message to the log;
%else %do;
        %if &event. = torn %then %do;
              %put WARNING: ->>> NO TORNADO HAS OCCURRED ON &date. <<<-;
        %end;
        %else %do;
              %put WARNING: ->>> NO %upcase(&event.) HAS OCCURRED ON
              &date. <<<-;
        %end;
%end;
%end;
```

Step 7. Close and invoke the macro call.

```
%mend;

** Invoke the macro to import the NOAA data;
%noaa(&start_date.,&end_date.);
```

Results

The data from the NOAA files can be used for various purposes, including reports, maps, presentations and portfolio analysis.

In the example below tornado touchdowns for the period August 2009 – June 2010 were mapped using the GIS data (Latitude and Longitude) from the files. This data can be overlaid on top of GIS information about individual properties in a bank's portfolio, which becomes a powerful tool for analysis. The combined information can be used to estimate potential losses from storms, evaluate needs for insurance adjusters, management presentations, etc.

Reported Tornado Activity
August 2009 - June 2010
Preliminary Data from the National Weather Service Center
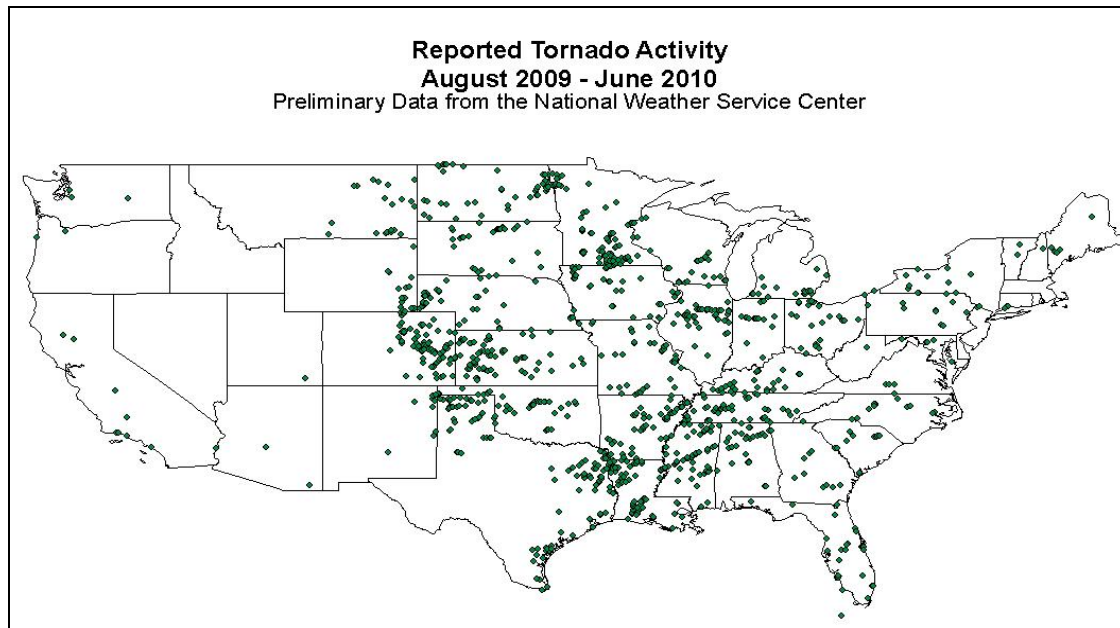
References

Conway, T. (2003) *The URL-y Show: Using SAS LE and the URL Access Method to Retrieve Stock Quotes*. SUGI Paper 73-28.
DeGuire, Y. (2007) *The FILENAME Statement Revisited*, Statistics Canada. SAS Global Forum Paper 007-2007.
SAS Institute Inc. 2010. *SAS 9.2 Language Reference: Dictionary*, Third Edition. Cary, NC

Acknowledgements

I am grateful to my wife Krasi Shtiliyanova for her unending patience to read my papers and analyses, and for the valuable comments, which has improved my writing significantly. I would like to thank my colleague Molly Israel for creating the map with the tornado activity.

Contact Information

Your comments and questions are valued and encouraged. You may contact the author at:

Valentin Todorov
Assurant Specialty Property
260 Interstate North Cir SE
Atlanta, GA 30339
Phone: 770-834-5161
valentin.todorov@assurant.com
www.vatodorov.com