Global Health Applications

Creating Multilingual Reports with the SAS Unicode Server

Sharon Coleman MS, MPH

December 10, 2015

Introduction

- Global public health initiatives are critical to improving the health of populations worldwide
- Research that incorporates data from various countries with different languages and scripts can be challenging
- The SAS Unicode server allows encoding of characters from multiple languages
- Bridging the language divide in research

What is the Unicode Transformation Format?

- A SAS session that uses UTF-8 : universal coding
- ASCII (American Standard Code for Information Interchange) codes represent text in computers and was the most common character encoding until 2007. Then UTF-8 surpassed ASCII
- ASCII only allows English and common punctuation but isn't useful for non-Latin characters

  128 ASCII characters vs. over a million code points!
- Code points are used with UTF-8 (**10101100** for "**€**")

Installing the Unicode server

- On Windows, you will find the additional language option on the SAS programs menu

Programs ► ► SAS

►SAS 9.3 (English)

► SAS 9.3 (Additional Languages)

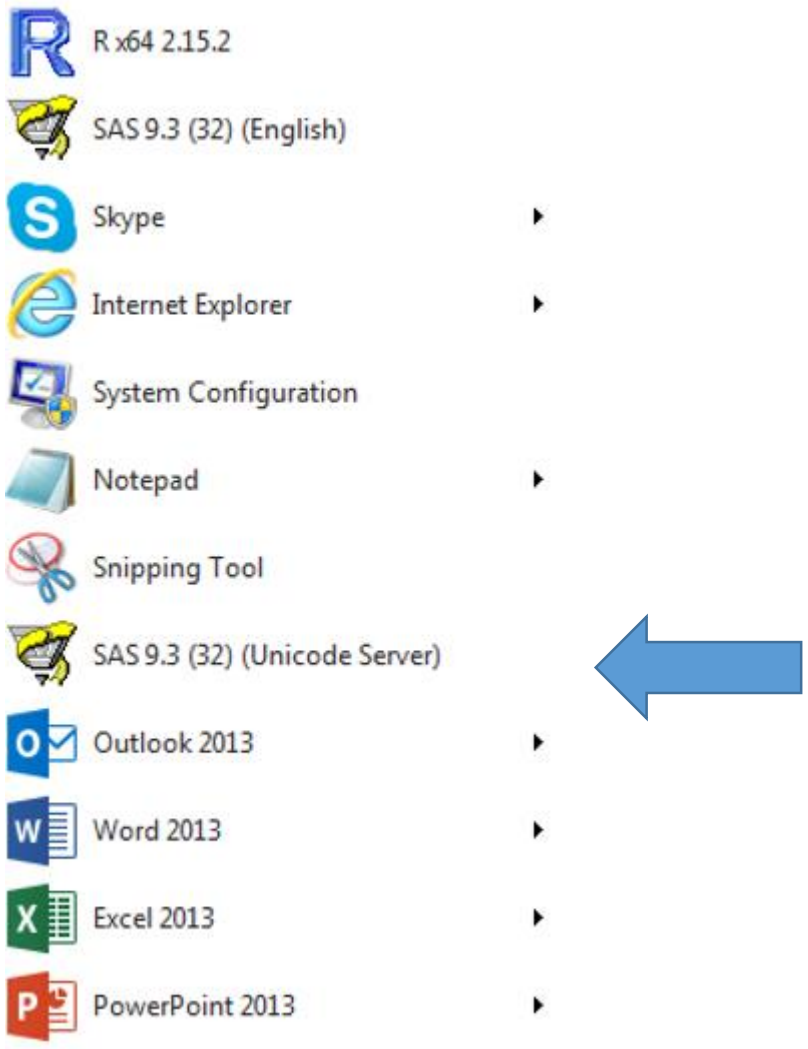
► SAS 9.3 (Unicode support)

*Please see contact SAS support for specific questions
http://support.sas.com/techsup/technote/ts801.pdf (p 6-7)

Multiple languages can be used in one session

	name	first	street	zip	city	country
60	Côté	Frédéric	2, rue descente...	36000	Châteauroux	France
61	Boucher	Corinne	21 place du Pant...	75009	Paris	France
62	Fournier	Étienne	91, rue Victor Hu...	71000	Mâcon	France
63	Cotée	Madeleine	691, avenue Fré...	30034	Nîmes	France
64	Legrand	Claire	62 rue de la Bret...	45000	Orléans	France
65	Dubois	Benoît	40 Rue Charles...	91330	Yerres	France
66	Thibeault	Georges	47 rue Alexandre...	91000	Évry	France
67	Martin	Désirée	90, rue Grisolle	83600	Fréjus	France
68	Vaudron	Sébastien	2 promenade de...	13008	Marseille	France
69	Girard	Régine	4 bis rue Profess...	69008	Lyon	France
70	Yılmaz	Ekrem	Tepebaşı Bulvarı...	80050	İstanbul	Türkiye
71	Kaya	Erkan	Akay Caddesi N...	06640	Ankara	Türkiye
72	Demir	Filiz	Fethi Cadessi No...	33212	İzmir	Türkiye
73	Şahin	Ali	Sanayi Cad. No:...	23300	Elazığ	Türkiye
74	Çelik	Nur	Kazımiye Mah. O...	59860	Çorlu	Türkiye
75	佐藤	明子	東京都渋谷区...	〒150-000	東京	日本
76	鈴木	幹夫	神奈川県横浜...	〒221-083	横浜市	日本
77	田中	あずみ	大阪府大阪市...	〒534-001	大阪市	日本
78	山本	麻美	厚木市寿町3-7-...	〒243-000	厚木市	日本

Start a SAS session that uses the UTF-8



Goal: Create a report of follow up study visits

Colleagues are in St. Petersburg Russia

Assistance needed in seeing follow up visits to implement study reminder calls

Example: Writing a report using Cyrillic script

subjectid	firstname	lastname	hphone1	cphone1	visit	scheddated
1002	Олег	Чаплыгина	8-812-750-62-66	8-812-750-62-66	6 месяцев	3/12/2015
1004	Роман	Викторович	8-812-750-62-66	8-812-755-66-66	7 месяцев	3/12/2015
1005	Алексей	Намик-оглы	8-812-750-62-66	8-813-500-62-66	8 месяцев	3/12/2015

```
data russia;
set work.russia;
format scheddated ddmmyy8.;
run;
```

Example: Writing a report using Cyrillic script

```
ods escapechar = '^' ;
options orientation = landscape nodate nonumber ;
ods noproctitle;
ods pdf file="C:\basug\scheduled_visits.pdf"
startpage=on startpage=now;
title1 bold height=11pt "Россия:Назначенные Отдаленные Интервью";
title2 bold height=9pt "Scheduled Follow-Up Visits";
title3 bold height=10pt " &dtnull3. ";
    proc report data=russia  nowd headline headskip box STYLE(header)=[font_weight=bold background=CXAEADD9
    vjust = center] missing;


    column subjectid firstname lastname hphone1 cphone1 visit scheddate ;

    define subjectid/center     "Номер участника"       style=[cellwidth=0.7in];
    define firstname/left       "Имя"                   style=[cellwidth=1.2in];
    define lastname/left        "Фамилия"               style=[cellwidth=1.2in];
    define hphone1/left         "Домашний Телефон"      style=[cellwidth=1.2in];
    define cphone1/left         "Мобильный Телефон"     style=[cellwidth=1.2in];
    define visit/left           "Вид Визита"            style=[cellwidth=0.8in];
    define scheddate/left       "Назначенная дата"      style=[cellwidth=0.8in];


    run;

ods pdf close;
```

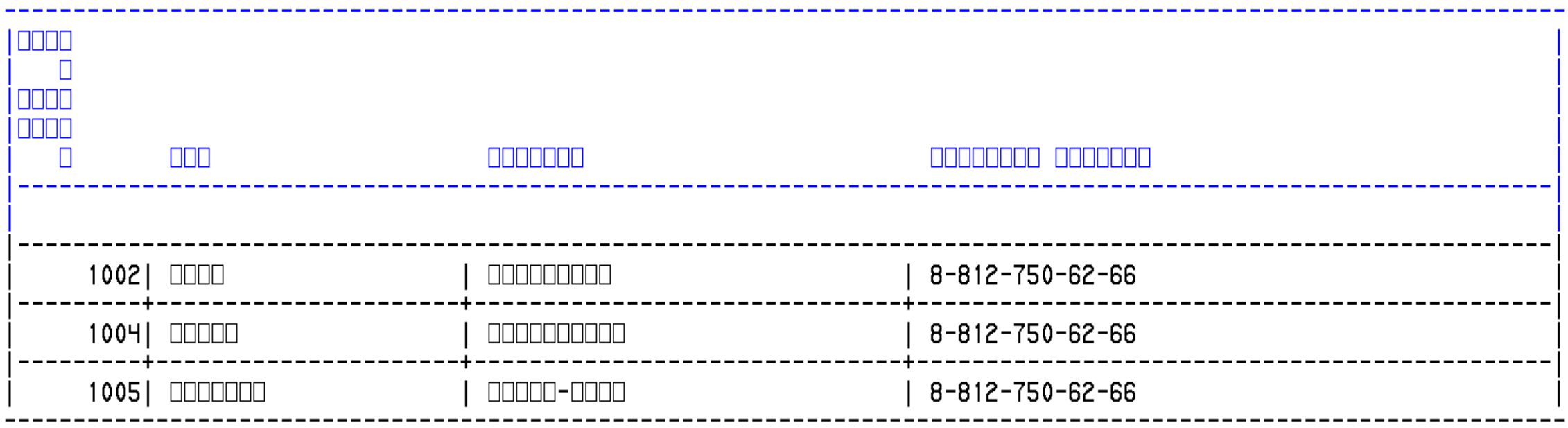
Warning!

- When you open SAS with Unicode Support, you will see the following in the log

WARNING: Display of UTF8 encoded data is not fully supported by the SAS Display Manager System.

You could use SAS Enterprise Guide or use an ODS output option

Output in Windows Display Manager



Report that is translated for the Russian team

Россия:Назначенные Отдаленные Интервью
Scheduled Follow-Up Visits
3 December 2015

Номер участника	Имя	Фамилия	Домашний Телефон	Мобильный Телефон	Вид Визита	Назначенная дата
1002	Олег	Чаплыгина	8-812-750-62-66	8-812-750-62-66	6 месяцев	03/12/15
1004	Роман	Викторович	8-812-750-62-66	8-812-755-66-66	7 месяцев	03/12/15
1005	Алексей	Намик-оглы	8-812-750-62-66	8-813-500-62-66	8 месяцев	03/12/15

Using National Language Support (NLS)

- Refer to —Values for the LOCALE= System Option SAS

http://support.sas.com/documentation/cdl/en/nlsref/63072/PDF/default/nlsref.pdf

Locale sensitive formatting

```
option locale=ru_RU;
data russia2;
set work.russia2;
format  Compensation  NLMNYI.;
run;
proc print;var subjectid lastname Compensation;
run;
```

Currency specific to the Country

Наблюдения	subjectid	lastname	Compensation
1	1002	Чаплыгина	1 000RUB
2	1004	Викторович	1 500RUB
3	1005	Намик-оглы	2 000RUB

Goal: Create a report of Chinese SMS messages

Colleagues are in China

Patient navigators will text a message in the participants preferred language

Data from an Excel spreadsheet

Subjectid	cellphone	SMS Text
1001	0839947259	安排复诊时间
1002	0839947259	检查结果
1003	0839947259	检查结果

```
proc import out= work.china
datafile= "C:\basug\china_report.xlsx"
dbms=excel replace;
getnames=yes;
mixed=no;
scantext=yes;
usedate=yes;
scantime=yes;
run;
```

```
ods escapechar = '^' ;
options orientation = landscape nodate nonumber ;
ods noproctitle;
ods rtf file="C:\basug\china_sms.doc"
startpage=on startpage=now;
title1 bold height=11pt "Prefered Language is Chinese";
title2 bold height=9pt "SMS Message";
title3 bold height=10pt " 3 December 2015 ";
footnote1 " Schedule Medical f/u visit";
footnote2 "check lab results";
    proc report data=china  nowd headline headskip box STYLE(header)=[font_weight=bold background=CXAEADD
    vjust = center] missing;
    column subjectid cellphone SMS_Text ;

    define subjectid/center     "SubjectID"                 style=[cellwidth=0.8in];
    define cellphone/left       "CellPhone"                 style=[cellwidth=1.2in];
    define SMS_Text/left        "SMS Text"                  style=[cellwidth=1.5in];
    run;
ods rtf close;
```

SMS text to be sent in Chinese characters

Prefered Language is Chinese
SMS Message
3 December 2015

SubjectID	CellPhone	SMS Text
1001	0839947259	安排复诊时间
1002	0839947259	检查结果
1003	0839947259	检查结果

Reference

- An excellent reference :

For more information about the concept of the SAS Unicode Server, see the technical Paper, —Processing Multilingual Data with the SAS® 9.2 Unicode Server.

Conclusion

- With the SAS Unicode server , you can meet global public health and research needs in a single SAS session!

Thank you: спасибо

Sharon Coleman

sharcole@bu.edu

File Management Using Pipes and 'X' Commands in SAS®

BU CTSI Biostatistics, Epidemiology & Research Design (BERD) Core Seminar

Emily K.Q. Sisson

December 10, 2013

Introduction

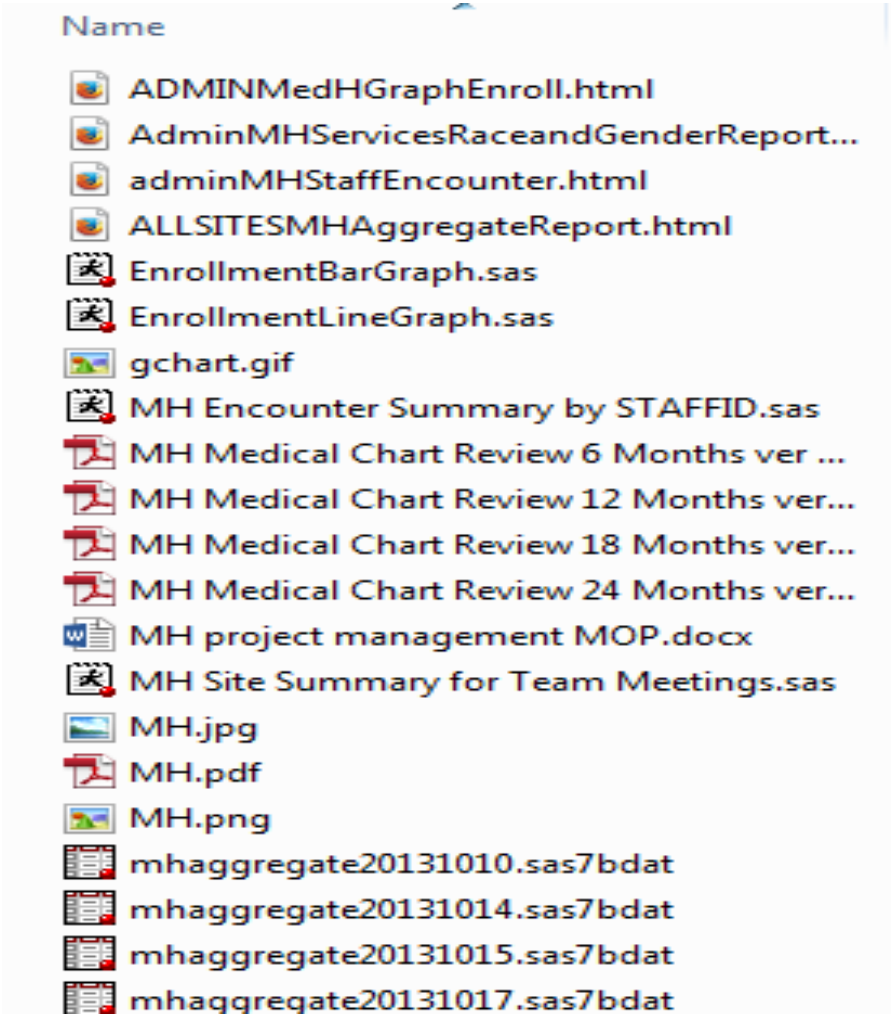
- We all know that SAS is a powerful analytic tool

- SAS also can be a great file management tool through the use of
  - Pipes
  - Operating System Commands (‘x’ commands)
  - Data manipulation
  - Call Execute Statements

Introduction

- Example problem: Hundreds of files of varying type in a single folder. As the project continues, it becomes increasingly difficult to locate documents

- For example, let's move these files into subdirectories within the current folder:
  - File Type (.html, .sas, etc)
    - Date (YYYYMM)



Introduction

- Generate a listing of files (pipe device)
- Utilize listing to manage and organize (data manipulation)
- Create new directories through SAS to archive your files (operating system (‘x’) commands; call execute)
- Move original files to new destinations (operating system (‘x’) commands; call execute)

Step 1: Generate a Database of Files

```
filename dirlist pipe 'dir "Y:\MyFiles" /s';

data dirlist ;
     length lineinfo $256 ;
     infile dirlist length=reclen ;
     input lineinfo $varying256. reclen ;
run;
```

Step 1: Generate a Database of Files

- Excerpt from DIRLIST dataset

	lineinfo
1	Volume in drive Y is Y03 - Y Drive (SPH)
2	Volume Serial Number is E65B-E2D5
3	
4	Directory of Y:\MyFiles
5	
6	09/28/2015 11:07 AM <DIR> .
7	09/28/2015 11:07 AM <DIR> ..
8	04/08/2014 10:33 AM 20,560 ADMINMedHGraphEnroll.html
9	03/20/2014 09:19 AM 67,438 AdminMHServicesRaceandGenderReport.html
10	09/28/2015 08:21 AM 799,745 adminMHStaffEncounter.html
11	09/28/2015 08:17 AM 197,282 ALLSITESMHAggregateReport.html
12	07/24/2015 08:58 AM 31,728 EnrollmentBarGraph.sas

Step 2: Utilize Listing to Manage and Organize

- Decide what it is, exactly, you want to do with your files
- Organize and archive by year? month? quarter?
- Delete large files?
- Develop new variables to help you organize the file list:
  - NEWDATE: *YYYYMM* of the file date
  - FILESIZE: Extract substring of file size from *LINEINFO*
  - PROGRAMSUITE: SAS, Office, SPSS, etc

Step 2: Utilize Listing to Manage and Organize

```
data dirlist_useful;
    set dirlist;
    /*The path of the directory only appears once, using a retain statement to assign*/
    length directory $1000;
    retain directory;
    if left(upcase(lineinfo))=:'DIRECTORY OF' then directory=substr(left(lineinfo),14);

    /*Isolate Other Important information*/
    filename = substr(left(lineinfo),40);              /*Name of file inclu extension*/
    fileextens = scan(strip(lineinfo),-1);                          /*File Type*/
    filedate = input(substr(lineinfo,1,10),?? mmddyy10.);          /*MM/DD/YYYY */
       format filedate mmddyy10.;
    filetime = input(scan(lineinfo,4)||" "||scan(lineinfo,5),time12.); /*HH:MM AM/PM*/
       format filetime time12.;
```

Step 2: Utilize Listing to Manage and Organize

```
    if month(filedate) < 10 then month =
        put(year(filedate),4.)||"0"||put(month(filedate),1.);      /*Character YYYYMM*/
    if month(filedate) ge 10 then month =
        put(year(filedate),4.)||put(month(filedate),2.);

    /*Delete extraneous rows*/
    if lineinfo = '' then delete;
    if index(upcase(lineinfo),'<DIR>') then delete;
    if left(upcase(lineinfo)) =: 'VOLUME' then delete;
    if left(upcase(lineinfo))=: 'DIRECTORY OF' then delete;
    if fileextens in ('bytes' 'Listed:' 'free') then delete;
run;
```

Step 2: Utilize Listing to Manage and Organize

- Excerpt from DIRLIST_USEFUL dataset

	directory	filename	fileextens	filedate	filetime	month
1	Y:\MyFiles	ADMINMedHGraphEnroll.html	html	04/08/2014	10:33:00	201404
2	Y:\MyFiles	AdminMHServicesRaceandGenderReport.html	html	03/20/2014	9:19:00	201403
3	Y:\MyFiles	adminMHStaffEncounter.html	html	09/28/2015	8:21:00	201509
4	Y:\MyFiles	ALLSITESMHAggregateReport.html	html	09/28/2015	8:17:00	201509
5	Y:\MyFiles	EnrollmentBarGraph.sas	sas	07/24/2015	8:58:00	201507
6	Y:\MyFiles	EnrollmentLineGraph.sas	sas	06/26/2015	13:19:00	201506
7	Y:\MyFiles	gchart.gif	gif	04/08/2014	4:04:00	201404
8	Y:\MyFiles	MH Encounter Summary by STAFFID.sas	sas	06/26/2015	13:19:00	201506
9	Y:\MyFiles	MH Medical Chart Review 12 Months ver 20150626.pdf	pdf	07/01/2015	9:14:00	201507
10	Y:\MyFiles	MH Medical Chart Review 18 Months ver 20150626.pdf	pdf	07/01/2015	9:14:00	201507
11	Y:\MyFiles	MH Medical Chart Review 24 Months ver 20150626.pdf	pdf	07/14/2015	13:11:00	201507

Step 3: Create New Directories

- Operating System (‘x’) commands can be used directly in a SAS session
- To suppress the RUN window, use the system option noxwait in your code
- Commands we’ll consider today:
  - mkdir – makes a directory (option –p will avoid errors on preexisting dir)
  - move – relocates a file from one directory to another

Step 3: Create New Directories

```
proc sql;
    create table dir_typ_month as
    select distinct directory, fileextens, month
    from dirlist_useful;
quit;

%macro createdir(dir=,typ=,mon=);
    x "mkdir -p  &dir.\&typ.\&mon.";
%mend createdir;

options noxwait;
data _null_;
    set dir_typ_month;
    command = cats('%createdir(dir=',directory,', typ=',
                   upcase(fileextens),', mon=',month,');');
    call execute(command);
run;
```

Step 3: Create New Directories

- Excerpted view of data _null_

	directory	fileextens	month	command
1	Y:\MyFiles	docx	201403	%createdir(dir=Y:\MyFiles, typ=DOCX, mon=201403);
2	Y:\MyFiles	gif	201309	%createdir(dir=Y:\MyFiles, typ=GIF, mon=201309);
3	Y:\MyFiles	gif	201404	%createdir(dir=Y:\MyFiles, typ=GIF, mon=201404);
4	Y:\MyFiles	html	201403	%createdir(dir=Y:\MyFiles, typ=HTML, mon=201403);
5	Y:\MyFiles	html	201404	%createdir(dir=Y:\MyFiles, typ=HTML, mon=201404);
6	Y:\MyFiles	html	201409	%createdir(dir=Y:\MyFiles, typ=HTML, mon=201409);
7	Y:\MyFiles	html	201509	%createdir(dir=Y:\MyFiles, typ=HTML, mon=201509);
8	Y:\MyFiles	jpg	201307	%createdir(dir=Y:\MyFiles, typ=JPG, mon=201307);
9	Y:\MyFiles	pdf	201310	%createdir(dir=Y:\MyFiles, typ=PDF, mon=201310);
10	Y:\MyFiles	pdf	201507	%createdir(dir=Y:\MyFiles, typ=PDF, mon=201507);
11	Y:\MyFiles	png	201307	%createdir(dir=Y:\MyFiles, typ=PNG, mon=201307);
12	Y:\MyFiles	pptx	201303	%createdir(dir=Y:\MyFiles, typ=PPTX, mon=201303);
13	Y:\MyFiles	sas	201501	%createdir(dir=Y:\MyFiles, typ=SAS, mon=201501);
14	Y:\MyFiles	sas	201505	%createdir(dir=Y:\MyFiles, typ=SAS, mon=201505);
15	Y:\MyFiles	sas	201506	%createdir(dir=Y:\MyFiles, typ=SAS, mon=201506);
16	Y:\MyFiles	sas	201507	%createdir(dir=Y:\MyFiles, typ=SAS, mon=201507);
17	Y:\MyFiles	sas	201508	%createdir(dir=Y:\MyFiles, typ=SAS, mon=201508);
18	Y:\MyFiles	sas7bdat	201310	%createdir(dir=Y:\MyFiles, typ=SAS7BDAT, mon=201310);
19	Y:\MyFiles	sas7bdat	201311	%createdir(dir=Y:\MyFiles, typ=SAS7BDAT, mon=201311);

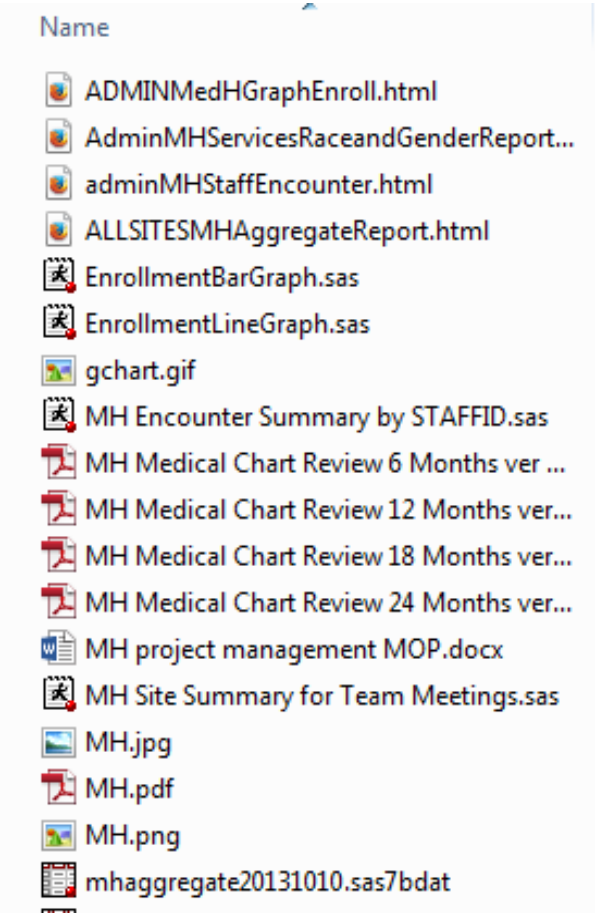
Step 4: Move Files to New Folder Structure

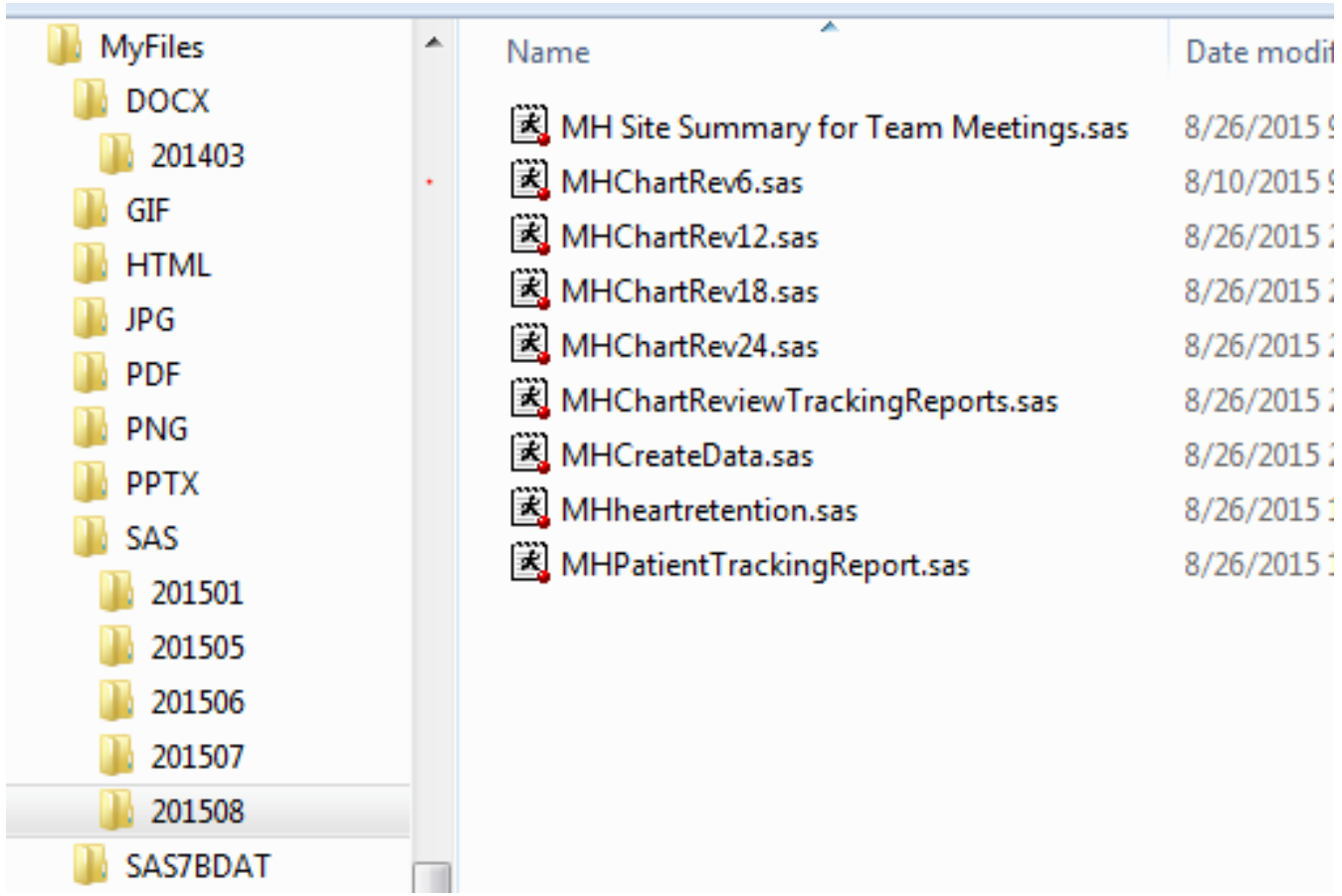
```
%macro movfil(dir=,name=,typ=,mon=);
    x move "&dir\&name" "&dir\&typ\&mon\&name";
%mend movfil;

data _null_;
    set dirlist_useful;
    command = cats('%movfil(dir=',directory,', name=',filename,
                   ', typ=',upcase(fileextens),', mon=',month,');');
    call execute(command);
run;
```

Step 4: Move Files to New Folder Structure

fileextens	filedate	filetime	month	command
html	04/08/2014	10:33:00	201404	%movfil(dir=Y:\MyFiles, name=ADMINMedHGraphEnroll.html, typ=HTML, mon=201404);
html	03/20/2014	9:19:00	201403	%movfil(dir=Y:\MyFiles, name=AdminMHServicesRaceandGenderReport.html, typ=HTML, mon=201403);
html	09/28/2015	8:21:00	201509	%movfil(dir=Y:\MyFiles, name=adminMHStaffEncounter.html, typ=HTML, mon=201509);
html	09/28/2015	8:17:00	201509	%movfil(dir=Y:\MyFiles, name=ALLSITESMHAggregateReport.html, typ=HTML, mon=201509);
sas	07/24/2015	8:58:00	201507	%movfil(dir=Y:\MyFiles, name=EnrollmentBarGraph.sas, typ=SAS, mon=201507);
sas	06/26/2015	13:19:00	201506	%movfil(dir=Y:\MyFiles, name=EnrollmentLineGraph.sas, typ=SAS, mon=201506);
gif	04/08/2014	4:04:00	201404	%movfil(dir=Y:\MyFiles, name=gchart.gif, typ=GIF, mon=201404);
sas	06/26/2015	13:19:00	201506	%movfil(dir=Y:\MyFiles, name=MH Encounter Summary by STAFFID.sas, typ=SAS, mon=201506);
pdf	07/01/2015	9:14:00	201507	%movfil(dir=Y:\MyFiles, name=MH Medical Chart Review 12 Months ver 20150626.pdf, typ=PDF, mon=201507);
pdf	07/01/2015	9:14:00	201507	%movfil(dir=Y:\MyFiles, name=MH Medical Chart Review 18 Months ver 20150626.pdf, typ=PDF, mon=201507);

Step 4: Move Files to New Folder Structure





Conclusion

- Pipe devices and operating system commands are powerful ways to communicate with your network
- With thoughtful use of the call execute function, hundreds, if not thousands, of files can be manipulated in a matter of seconds
- Archiving files is just one of many possibilities for these powerful tools

References

- Creating a Directory Listing Using SAS for Windows
  - http://support.sas.com/kb/24/820.html

- Running Windows or MS-DOS Commands from within SAS
  - https://support.sas.com/documentation/cdl/en/hostwin/63285/HTML/default/viewer.htm#exittemp.htm
  - http://www2.sas.com/proceedings/forum2008/092-2008.pdf

- CALL EXECUTE Routine
  - http://support.sas.com/documentation/cdl/en/mcrolref/61885/HTML/default/viewer.htm#a000543697.htm

Contact

Emily K.Q. Sisson

Boston University School of Public Health

Data Coordinating Center

801 Massachusetts Avenue, 3rd Floor

Boston, Massachusetts 02118

(617) 638-5869

eq@bu.edu

Comments and Questions are valued and encouraged!