

SAS/AF Frame in programmer toolkit

Patrick Lupien, Principal Statistical Programmer
Berlin, October 19, 2010

Agenda

- Why SAS/AF in a toolkit?
- Data compare
- Dataset viewer
- Metadata viewer
- Conclusion
- Questions

Why SAS/AF in a toolkit?

- Ease common programming tasks
- Provide GUI for non SAS user

Data compare

- It is often useful to browse 2 datasets at the same time

VIEWTABLE: US Zipcodes; Source: zipcodedownload.com April 2008, SAS Release 9.2

	ZIP	CITY	STATE	STATECODE	STATENAME	COUNTY	COUNTYNM
1	00501	Holtsville	36	NY	New York	103	Suffolk
2	00544	Holtsville	36	NY	New York	103	Suffolk
3	00601	Adjuntas	72	PR	Puerto Rico	1	Adjuntas
4	00602	Aguada	72	PR	Puerto Rico	3	Aguada
5	00603	Aguadilla	72	PR	Puerto Rico	5	Aguadilla
6	00604	Aguadilla	72	PR	Puerto Rico	5	Aguadilla
7	00605	Aguadilla	72	PR	Puerto Rico	5	Aguadilla
8	00606	Maricao	72	PR	Puerto Rico	93	Maricao

VIEWTABLE: Work.Zipcode

	ZIP	CITY	STATE	STATECODE	STATENAME	COUNTY	COUNTYNM
1	00501	Holtsville	36	NY	New York	103	Suffolk
2	00544	Holtscity	36	NY	New York	103	Suffolk
3	00601	Adjuntas	72	PR	Puerto Rico	1	Adjuntas
4	00602	Aguada	72	PR	Puerto Rico	3	Aguada
5	00603	Aguadilla	72	PR	Puerto Rico	5	Aguadilla
6	00604	Aguadilla	72	PR	Puerto Rico	5	Aguadilla
7	00606	Maricao	72	PR	Puerto Rico	93	Maricao
8	00610	Anasco	72	PR	Puerto Rico	11	Anasco

- Both datasets are independant
Column order can differs
Where clause need to be apply to both
Datasets are locked

Data compare

PL Tools Dataset Compare Version 1 Build 08

Base

Libname: Sashelp ZIPCODE Number of obs: 41761 Use version: Version

	ZIP	Y	X	ZIP_CLASS	CITY	STATE	STATECODE	STATENAME	COUNTY
1	00501	40.813078	-73.046388	U	Holtsville	36 NY	New York	103	
2	00544	40.813223	-73.049288	U	Holtsville	36 NY	New York	103	
3	00601	18.165950	-66.723627		Adjuntas	72 PR	Puerto Rico	1	
4	00602	18.383005	-67.186553		Aguada	72 PR	Puerto Rico	3	
5	00603	18.433236	-67.151954		Aguadilla	72 PR	Puerto Rico	5	
6	00604	18.505289	-67.135899	P	Aguadilla	72 PR	Puerto Rico	5	
7	00605	18.436149	-67.151346	P	Aguadilla	72 PR	Puerto Rico	5	
8	00606	18.185616	-66.977377		Maricao	72 PR	Puerto Rico	93	
9	00610	18.285948	-67.144161		Anasco	72 PR	Puerto Rico	11	
10	00611	18.287716	-66.797578	P	Angeles	72 PR	Puerto Rico	141	
11	00612	18.471326	-66.728149		Arecibo	72 PR	Puerto Rico	13	
12	00613	18.472737	-66.719283	P	Arecibo	72 PR	Puerto Rico	13	

Compare

Libname: Work ZIPCODE Number of obs: 41760 Use version: Version

	ZIP	Y	X	ZIP_CLASS	CITY	STATE	STATECODE	STATENAME	COUNTY
1	00501	40.813078	-73.046388	U	Holtsville	36 NY	New York	103	
2	00544	40.813223	-73.049288	U	Holtsville	36 NY	New York	103	
3	00601	18.165950	-66.723627		Adjuntas	72 PR	Puerto Rico	1	
4	00602	18.383005	-67.186553		Aguada	72 PR	Puerto Rico	3	
5	00603	18.433236	-67.151954		Aguadilla	72 PR	Puerto Rico	5	
6	00604	18.505289	-67.135899	P	Aguadilla	72 PR	Puerto Rico	5	
7	00605			P					
8	00606	18.185616	-66.977377		Maricao	72 PR	Puerto Rico	93	
9	00610	18.285948	-67.144161		Anasco	72 PR	Puerto Rico	11	
10	00611	18.287716	-66.797578	P	Angeles	72 PR	Puerto Rico	141	
11	00612	18.471326	-66.728149		Arecibo	72 PR	Puerto Rico	13	
12	00613	18.472737	-66.719283	P	Arecibo	72 PR	Puerto Rico	13	

Find OBS: Find Where: Apply Hide/Unhide Compare Refresh OK

Data compare

How to synchronize vertically

- Set the vertical scrollbar maximum property to the maximum obs number between both dataset
- Set the vertical scrollbar property largeScrollAmount to the number of obs visible in the viewer
- On vertical scroll, both viewer _setTopRow property are set base on the vertical scrollbar value
- Ensure selected obs is not bigger than the last obs in one of the viewer

Data compare

How to synchronize columns and subset

- Default column order inherited from top dataset.
- Columns found only in bottom dataset are displayed on the right
- Horizontal scroll follow same rule but using `setLeftColumn` property for each viewers
- Subset or direct move to an obs number apply to both viewers

Data compare

Basic features

- Displayed column in both viewers keep together
- Horizontal and vertical scrollbars are synchronizing both viewers
- Where condition or move to an observation number performed on both viewers
- Viewers display copy of the related datasets
Thus avoiding locking issues
Display can be refresh while keeping current layout and subset

Data compare

More advanced features

- Traffic lightning base on results of proc compare
Yellow for obs with a difference
Red for column with a difference
- Provided a dataset is in source control, allow display of any versions
- Allow screen resizing

Dataset viewer

- Same features as Data compare
- One viewer to display more rows
- Main advantage is no dataset lock

Dataset viewer

PL Tools Dataset Viewer Prototype Version 1 Build 01b

Libname: Sashelp Sashelp.ZIPCODE Number of obs: 41761 Use version: Version

	ZIP	Y	X	ZIP_CLASS	CITY	STATE	STATECODE	STATENAME	COUNTY	COU
1	00501	40.813078	-73.046388	U	Holtsville	36	NY	New York	103	Suffo
2	00544	40.813223	-73.049288	U	Holtsville	36	NY	New York	103	Suffo
3	00601	18.165950	-66.723627		Adjuntas	72	PR	Puerto Rico	1	Adju
4	00602	18.383005	-67.186553		Aguada	72	PR	Puerto Rico	3	Agua
5	00603	18.433236	-67.151954		Aguadilla	72	PR	Puerto Rico	5	Agua
6	00604	18.505289	-67.135899	P	Aguadilla	72	PR	Puerto Rico	5	Agua
7	00605	18.436149	-67.151346	P	Aguadilla	72	PR	Puerto Rico	5	Agua
8	00606	18.185616	-66.977377		Maricao	72	PR	Puerto Rico	93	Maric
9	00610	18.285948	-67.144161		Anasco	72	PR	Puerto Rico	11	Anas
10	00611	18.287716	-66.797578	P	Angeles	72	PR	Puerto Rico	141	Utua
11	00612	18.471326	-66.728149		Arecibo	72	PR	Puerto Rico	13	Areci
12	00613	18.472737	-66.719283	P	Arecibo	72	PR	Puerto Rico	13	Areci
13	00614	18.456904	-66.735892	P	Arecibo	72	PR	Puerto Rico	13	Areci
14	00616	18.426456	-66.673779		Bajadero	72	PR	Puerto Rico	13	Areci
15	00617	18.458792	-66.537234		Barceloneta	72	PR	Puerto Rico	17	Barc
16	00622	18.017976	-67.170627		Boqueron	72	PR	Puerto Rico	23	Cabo
17	00623	18.086721	-67.151780		Cabo Rojo	72	PR	Puerto Rico	23	Cabo
18	00624	18.063481	-66.723612		Penuelas	72	PR	Puerto Rico	111	Penu
19	00627	18.485866	-66.844057		Camuy	72	PR	Puerto Rico	27	Cam
20	00631	18.180272	-66.832917	P	Castaner	72	PR	Puerto Rico	81	Lares
21	00636	18.162126	-67.077928	P	Rosario	72	PR	Puerto Rico	125	San
22	00637	18.081674	-66.968022		Sabana Grande	72	PR	Puerto Rico	121	Saba
23	00638	18.338774	-66.466418		Ciales	72	PR	Puerto Rico	39	Ciale
24	00641	18.268891	-66.704425		Utuado	72	PR	Puerto Rico	141	Utua
25	00646	18.466563	-66.270521		Dorado	72	PR	Puerto Rico	51	Dora
26	00647	17.967309	-66.939935		Ensenada	72	PR	Puerto Rico	55	Guar
27	00650	18.362641	-66.565388		Florida	72	PR	Puerto Rico	54	Florid
28	00652	18.466248	-66.569371		Garrochales	72	PR	Puerto Rico	13	Areci
29	00653	17.974015	-66.908699		Guanica	72	PR	Puerto Rico	55	Guar
30	00656	18.025970	-66.787588		Guayanilla	72	PR	Puerto Rico	59	Guay

Find OBS: Find Where: Apply Hide/Unhide Refresh OK

Metadata viewer

- SAS dictionary table columns have metadata for all accessible columns
- Allow searching for information base on column name or column label between datasets and library

Metadata Viewer

PL Tools Metadata Viewer Version 1 Build 01a

Libname

MSTOOLS
MAPS
SASUSER
WORK
SASHELP

Read

Libname (Library Name)	Memname (Dataset Name)	Name (Column Name)	Label (Column Label)
WORK	COUNTRY	fmtname	
WORK	COUNTRY	label	GLC Country Name
WORK	COUNTRY	start	ISO Country Number: 900+ Undefined
WORK	DEMOGRAPHICS	AdolescentFPpct	Adolescent Fertility Proportion Percentage
WORK	DEMOGRAPHICS	AdolescentFPyear	Adolescent Fertility Proportion Year
WORK	DEMOGRAPHICS	AdultLiteracyPct	Adult Literacy Rate Percentage (2000-2004)
WORK	DEMOGRAPHICS	CONT	Numeric Rep. for Continent
WORK	DEMOGRAPHICS	FemaleSchoolpct	Net Primary School Enrollment Ratio - Female
WORK	DEMOGRAPHICS	GNI	Gross National Income per Capita (PPP Int.\$ 2
WORK	DEMOGRAPHICS	ID	GLC Country ID Number
WORK	DEMOGRAPHICS	ISO	ISO Country Number: 900+ Undefined
WORK	DEMOGRAPHICS	ISONAME	ISO Name for Country
WORK	DEMOGRAPHICS	MaleSchoolpct	Net Primary School Enrollment Ratio - Male Pe
WORK	DEMOGRAPHICS	PopPovertyYear	Population Living Below the Poverty Line (Yea
WORK	DEMOGRAPHICS	PopPovertyPct	Population Living Below the Poverty Line (% w
WORK	DEMOGRAPHICS	pop	Population (2005)
WORK	DEMOGRAPHICS	popAGR	Population Annual Growth Rate Percentage (
WORK	DEMOGRAPHICS	popUrban	Population in Urban Areas Percentage (2005)
WORK	DEMOGRAPHICS	region	Region
WORK	DEMOGRAPHICS	totalFR	Total Fertility Rate (per woman 2004)
WORK	ZIPCODE	ALIAS_CITY	Alternate names of city separated by
WORK	ZIPCODE	AREACODE	Single Area Code for ZIP Code.
WORK	ZIPCODE	AREACODES	Multiple Area Codes for ZIP Code.
WORK	ZIPCODE	CITY	Name of city/org
WORK	ZIPCODE	COUNTY	FIPS county code.
WORK	ZIPCODE	COUNTYNM	Name of county/parish.
WORK	ZIPCODE	DST	ZIP Code obeys Daylight Savings: Y:Yes N-N

Apply where Sort

Dataset WORK.DEMOGRAPHICS column ISO details

Format name : CTRY_.

[Start-End] Label

[4 - 4] AFGHANISTAN

[8 - 8] ALBANIA

[12 - 12] ALGERIA

[20 - 20] ANDORRA

[count] - Unique value

[1] AFGHANISTAN

[1] ALBANIA

[1] ALGERIA

[1] ANDORRA

[1] ANGOLA

[1] ANTIGUA/ BARBUDA

[1] AZERBAIJAN

[1] ARGENTINA

[8] AUSTRALIA

[1] AUSTRIA

[1] BAHAMAS

[1] BAHRAIN

[1] BANGLADESH

[1] ARMENIA

[1] BARBADOS

Statistics

Subset value

N : 197

NMiss : 0

NUnique : 192

Mean : 432.837563461776

Std : 269.625497174803

Median : LATVIA

Min/Max : AFGHANISTAN / ZAMBIA

Q1/Q3 : BENIN / ST. LUCIA

Apply where

OK

Metadata Viewer

Main features

- Display all columns for selected library
- Allow for subset and sort the columns
- Allow review of column contents by displaying
 - content of any user defined format attach to a column
 - unique values with a count for each
 - some basic statistics on the column content
 - subset of the statistics

Conclusion

- SAS/AF can be use to ease daily work and increase efficiencies
- Provide access to data for non SAS users

Questions ?

