EpiData Analysis

Specification of principle for command and programme control structures.

v1.0 – Oct. 2006 . Author: JM. Lauritsen, EpiData Association.

Default behaviour of software is bound to tradition and experience. The intention of EpiData software design is to allow for user control in a simple manner, but based on replication principles and good practices of data definitions (meta data) and the simplicity logic implemented in Stata (www.stata.com), where the philosophy is that the user should only indicate as much as needed, but also that the user is in complete control.

Approach to control in EpiData Analysis

Interface and execution of commands is controlled by implementation of a number of rules and definitions. There is a separation of what the user can control (run time defined) and what is built in (design time defined -hard coded) and only under control of the development team. The general rule is that overall logic and performance is defined as built in to ensure consistency, but for aspects typically dominated by user preference that part would be run-time rules. The typical user will not need to know all of this, but will see the usage when using assisted command building tools, e.g. the graph specification wizards.

Three types of **run time specification rules** have been implemented:

a. Rules that work in general – these are called ”**set specifications**”.

b. Rules that apply for a given command for a single ”run” of that command – these are called **options**.

c. Rules which define default setting of options.

a: Generally affect interface (menu's, character set, font, sizes) or formats (table design, format). Most of these will only be of interest to advanced users.

b: Control a given command. Eg. how many decimals in percentages in current table

c: An easy way to define default, e.g. if the user always wants 2 decimals in percentages. Technically these are implemented as ”set specicifations”.

How to see definitions and options ?

Set : Execute the command ”**set**”, which will show current value of all currently defined ”set”

Options: Are shown in the command reference (the file shown when pressing F1) with each command. Notice that some options are general and therefore not repeated with each command.

In a later phase EpiData Analysis will have an ”options” part in the file menu, untill then users must issue ”set specifications” as commands or includie in epidataanalysis.ini

The next sections of this documents shows current definitions.

Set Specification commands.

All must be used as : set < specification from table below > = value
e.g.: set display command history=off

		notes
Interface control	**Values**	
DISPLAY COMMAND HISTORY	on off	Depending on setting the user can specify
DISPLAY COMMAND PROMPT	on off	which of the interface elements should be
DISPLAY COMMANDTREE	on off	open.
DISPLAY DATABROWSER	on off	
DISPLAY MAINMENU	on off	
DISPLAY TOOLBAR	on off	
DISPLAY VARIABLES	on off	
DISPLAY WORKTOOLBAR	on off	
LANGUAGE	various	language of menu's, help files tc.
EDITOR PRINT INFO	on off	add date, time when printing from editor
PRINT PREVIEW CM	on off	work in centimeter (on) or inch(off)
Interface font size and type:	**(examples)**	Alfabeth type etc., e.g. for Chinese:
VIEWER FONT CHARSET	ISO-8859-1	SET viewer font charset = "gb2312"
VIEWER FONT NAME	Verdana, Times	set viewer font name="Arial Unicode MS"
VIEWER FONT SIZE	10	
WINDOW FONT SIZE	12	All supplementary parts
EDITOR FONT SIZE	12	
STYLE SHEET	filename	style sheet name (default: epiout.css)
STYLE SHEET EXTERNAL	on off	include styles in output or not
START PAGE	filename	file to show on startup (default: start.htm)
Information after commands:		
ECHO	on off	Defines what is going to be shown after
SHOW COMMAND	on off	execution of a command. E.g. show
SHOW ERROR	on off	info=on will result in all supplementary
SHOW INFO	on off	information.
SHOW RESULT	on off	echo is a short form for:
		result+command+info
Reading data		include (on) or exclude (off) records
READ DELETED	off on	marked for deletion
Graph definitons		
GRAPH CLIPBOARD	on off	always copy to clipboard
GRAPH COLOUR	default	color numbers for series, e.g. 11111111
GRAPH FILENAME FOLDER	on off	full folder name or no folder
GRAPH FILENAME SHOW	on off	include in footnote below graph
GRAPH FONT SIZE	10	
GRAPH FOOTNOTE	text	default is: ”EpiData Analysis Graph”
GRAPH SAVE	on off	save graph
GRAPH SAVETYPE	png wmf bmp	default: png
GRAPH SIZE X	400	
GRAPH SIZE Y	300	
Table design		
TABLE DESIGN	line	The style sheet class name definitions
TABLE DESIGN FREQ	line	
TABLE DESIGN GRAPH	graph	
TABLE DESIGN STAT	filled	

TABLE DESIGN SYSTEM	system	
Percent formats in tables TABLE PERCENT FORMAT COL TABLE PERCENT FORMAT ROW TABLE PERCENT FORMAT TOTAL TABLE PERCENT HEADER TABLE PERCENT HEADER COL TABLE PERCENT HEADER ROW TABLE PERCENT HEADER TOTAL	P1() P1{} P1[] % % % % %	
Format for confidence intervals TABLE CI FORMAT TABLE CI HEADER	 C()- (95% CI)	
History and default output HISTORY COMMAND PGM HISTORY COMMENT HISTORY NAME OUTPUT FOLDER OUTPUT NAME OUTPUT OPEN	 on off on off filename folder name eaoutput.htm on off	 Add to history when running pgm files Include comments in history Upon quit, save history file name (def: temp) Folder for output and history standard logfile name at initiation turn logfile on at start of programme
User preference for output: SHOW VAR VALUE SHOW VAR NAME	 default/V/VV default/VN/VNN	 showing of variable and value labels. see in ”options” for clarification
Variable creation RECODE INTERVAL TEXT VAR GENERATE TYPE	 - f/i	 text to put btw. x – y in recode values Default variable type with ”**gen**” command. f:float, i: integer
Internal development information DEBUG FILENAME DEBUG LEVEL SHOW SYSTEMINFO	 filename 0...5 ON OFF	For internal development. Writes out information at subroutine level depending on value of debug level

General Options		
Q	”Quiet” - Hide output for this command	/Q
W	Frequency Weighted counts on variable.	tab v1 /w=”wvar”
Default (no options)	Show variable label and value labels, but exclude var name if included in start of label	
V	Show value labels and values	/VV
VV	Show values instead of value labels	/V
VN	Show variable name and label	/VN
VNN	Only variable name, exclude label	/VNN

Options for frequencies, crosstabulation and statistics tables

Table cell contents and summary tables		
M	Allow missing variables as a category	/M
CD	Complete Data=no missing in any variable	/CD
S	Summary table shown after other tables. (Content of summary table defined by other options)	/S
NT	Do not show subtables (when stratifying)	/NT
NC	Do not show crude table when stratifying	/NC
OA	Outbreak analysis (first var=ill versus v_2...v_n)	/OA
MT	”Multi” Crosstables (First variable x all other)	/MT
MR	Multiple Response (not implemented yet) – (count several variables into one table and show number of responses and number of cases)	/MR
R C TO	Show Row Column Total percentages	/R /C /TO
PCT	Percentages in separate columns	/R /C /PCT
CP	Cumulative Percentages shown in (frequency tables)	
D0 D1 D2	Number of decimals in percentages	/R /D0

Sort Options in crosstables (and frequency tables)		
default (no options)	ascending order by values of column & Row	
SD	descending on value of 1st column and 1st row	/SD
SRAT SCAT SRDT SCDT	**{R**ow/**C**olumn}{**A**scending/**D**escending}**T**otal	/SRAT /SCAT
SLD SLA	ascending/descending on labels row+column	/SLA
SRA= SRD= SCA= SCD=	Specified column/row for sorting (numerical): **{R**ow/**C**olumn**}{A**scending/**D**escending**}T**otal	/SRA=2 /Scd=2

Estimation and statistical tests			
all tables	T	Test: Chi^2 for tables. F or T test in means	/T
..	GAM	Goodmann & Kruskall Gamma	/Gam
..	EX	Exact Test (Fishers Exact) for nxk tablesRelative	/ex
2x2 tables	RR	Risk (with /Epi replace Odds for RR)	/rr
..	ODS	Odds Ratio's	/ods
..	EPI	Short form for Odds Ratio + Chi^2	/epi
..	AR pAR	Attributable Risk Percent, Population Att.Risk	/AR /PAR
Frequency	CI	Confidence Interval (frequency tables)	/ci
All	E1 E2 E3 E4 E5	Decimal points for the estimators and statistics	/E3
summary of stratified	T	Mantel-Haenzel/Partial Gamma	/t
	OBS	In summary table for 2x2 all counts are shown	/obs
Exact tests	/sim	Number of simulations (default 500)	/sim=1000
	/seed	Seed for random number generator (default 9)	/Seed=25

Options all Graphs	explanation	example
hgrid vgrid	Show grid lines within graph area	/hgrid
xhide yhide	Do not show axis (x resepectively y axis)	/yhide
xinv yinv	Reverse scaling axis	/xinv
xlog ylog	Change axis scale to logaritmic	/ylog
noxtick noytick	Remove ticks on axis	/noxtick
xmin xmax ymin ymax	Force minimum and maximum to these values	/xmin=1 xmax=10
xtext ytext	Text to show as axis explanation	/xtext=”1,20,yes,1”
xline yline	Show line(-s) at value	/xline=10
xlined ylined	Show dotted line(-s) at value	/ylined=15
xinc yinc	Axis increment	/xinc=5
xl90	Change angle of texts on x axis to 90	/XL90
yvalue	Show Y value in a small box next to each (x,y)	/yvalue
save	Save graph to a file with indicated name	/save=”p3.png”
Options in SPC Charts		
B (was break)	Divide spc graph at breakpoint	/B
NT (was notest)	No tests are applied	/nt
NEGLCL	Allow negative lower control limit	/neglcl
EXP	Exclude point	/exp=10
EXV	Exclude values from this and above	/exv=12
EXZ	Exclude points where Y=0	/exz
Xlab	Use values from variable for ticks on X axis.	/xlab=v2
Tab	Show table of counts below graph	/tab
Other Special Graph Options		
xonly	Histogram: Only include x values with data	/xonly
xall	Bar: Include all x values from min to max	
sort	**H**istogram and **B**ar: sort on x value	
NO (was hideout)	Boxplot: Hide outliers = No outliers	/NO

Other Command specific options

command and option		***explanation***	***example***
Erasepng	all noconfirm	Erase all png files regardless of name Do so without confirmation for each file.	
read	close	close any open data file	
type	H1 H2 H3 H4 H5 CLASS	indicates the paragraph html code for the added text.	
show	class	class to use for text in a file shown with ”show”	/class=”p”
stattables aggregate	p1..p99 MEAN IQR IDR DES	(for a complete list see command reference) -	
LogOpen	A (append) Close Replace	append to existing file close an open logfile Replace the file indicated (overwrite)	/Close /replace
Agg Stattables	Close	Replace current data in memory by aggregated file	
List	NO	No observation numbers shown	/NO
file commands which save	REPLACE	For (savedata, savepgm, logopen, graphs – save) allow replacing of file with same name	/REPLACE