

EpiData Analysis



Specification of principle for command and programme control structures.

v1.2 – Nov. 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 specifications”**.

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

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
Exact tests RANDOM SEED RANDOM SIMULATIONS	9 500	Number of simulations and starting point for random seed.
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
	W	Frequency Weighted counts on variable.
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	Show values instead of value labels
	VN	Show variable name and label
	VNN	Only variable name, exclude label

Options for frequencies, crosstabulation and statistics tables

Table cell contents and summary tables		
	M	Allow missing variables as a category
	CC	Complete Cases=no missing in any variable
	S	Summary table shown after other tables. (Content of summary table defined by other options)
	NT	Do not show subtables (when stratifying)
	NC	Do not show crude table when stratifying
	OA	Outbreak analysis (first var=ill versus v ₂ ...v _n)
	MT	”Multi” Crosstables (First variable x all other)
	MR	Multiple Response (not implemented yet) – (count several variables into one table and show number of responses and number of cases)
	R C TO	Show <u>R</u> ow <u>C</u> olumn <u>T</u> otal percentages
	PCT	Percentages in separate columns
	CP	Cumulative Percentages shown in (frequency tables)
	D0 D1 D2	Number of decimals in percentages

Sort Options in crosstables (and frequency tables)

default (no options)		ascending order by values of column & Row
SD		descending on value of 1 st column and 1 st row
SRAT SCAT SRDT SCDT		{Row/Column} {Ascending/Descending} Total
SLD SLA		ascending/descending on labels row+column
SRA= SRD= SCA= SCD=		Specified column/row for sorting (numerical): S{Row/Column}{Ascending/Descending}

Estimation and statistical tests

all tables	T	Test: Chi ² for tables. F or T test in means
..	GAM	Goodmann & Kruskall Gamma
..	Ex	Exact Test (Fishers Exact) for nxk tables
2x2 tables	RRC	Risk Ratio (counts) + Chi ² (not implemented yet)
..	RRY	Risk Ratios (Population at risk) (not implemented yet)
..	O	Odds Ratio + Chi ²
..	OA	Outbreak Analysis Table
..	CI	Confidence Interval (frequency tables)
Frequency	S	Summary table of counts, statistics etc.
All	E1 E2 E3 E4 E5	Decimal points for the estimators and statistics
summary of stratified	T	Mantel-Haenzel/Partial Gamma
	OBS	In summary table all counts are shown (2x2 only)

Options all Graphs	explanation	example
hgrid vgrid xhide yhide xinv yinv xlog ylog noxtick noytick xmin xmax ymin ymax xtext ytext xline yline xlined ylined xinc yinc xl90 yvalue save edit ti sub fn	Show grid lines within graph area Do not show axis (x reasepectively y axis) Reverse scaling axis Change axis scale to logaritmic Remove ticks on axis Force minimum and maximum to these values Text to show as axis explanation Show line(-s) at value Show dotted line(-s) at value Axis increment Change angle of texts on x axis to 90 Show Y value in a small box next to each (x,y) Save graph to a file with indicated name Open graph for editing Title, subtitle and footnote texts	/hgrid /yhide /xinv /ylog /noxtick /xmin=1 xmax=10 /xtext="1,20,yes,1" /xline=10 /ylined=15 /xinc=5 /XL90 /yvalue /save="p3.png" /edit /fn="My project" /ti="title" /sub="sub title"
Options in SPC Charts		
B (was break) T NT (was notest) NEGLCL EXP EXV EXZ XI Tab	Divide spc graph at one or more breakpoints Apply testing Do not show table of tests Allow negative lower control limit Exclude point Exclude values from this and above Exclude points where Y=0 Use values from variable for ticks on X axis. Show table of counts below graph	/B=12 /b=24 /T /nt /neglcl /exp=10 /exv=12 /exz /xl=v2 /tab
Other Special Graph Options		
xonly xall sort	Histogram: Only include x values with data Bar: Include all x values from min to max Histogram and Bar: sort on x value	/xonly
NO (was hideout)	Boxplot: Hide outliers = No outliers	/NO

Other Command specific options

command and option		<i>explanation</i>	<i>example</i>
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