Sheet name	Example source scripts:												
	SPSS	Stata	SAS	R	Python								
source tables dataframes used in examples													
Sequential	Х	х											
Sequential with Replace			х										
Sequential with Update					x								
Sequential with Duplicate Column name					x								
1:1 merge	Х	х	х										
1:1 newRow	Х												
ful_join 1				х	x								
ful_join 2				Х	x								
full_join 3		х		Х	х								
left_join 1				Х	x								
left_join 2				Х	х								
right_join 1				Х	х								
right_join 2				Х	x								
inner join 1				х	х								
inner join 2				Х	х								
semi_join				Х									
semi_join 2				X									
unmatched				Х									
m:1		х											
Update 1		х	х										
Update 2		х	х										
SAS Update 1			х										
SAS Update 2			Х										
SAS Update Reverse 1			х										
SAS Update Reverse 2			х										
SAS Modify Reverse 1			х										
SAS Modify Reverse 2			х										
Replace 1			х										
Replace 2					х								
m:1 no newRow	х												
MergeFlagVariable	х												
MergeFlagVariable 2					х								
MergeFlagVariable 3		х											
MergeFlagVariable 4		х											
FirstVariable	Х												
Different Keys					x								
Different Keys 2					x								
Index Merge					x								

DSmerge1					DSmerge3			
caseID		firstName	income	famID	caseID	lastName	income	famID
	1	Abe	15000	11	1	Lincoln	_	
	2	Barry		11	2	Madison	32000	.a
	3	Chuck	17000	12	3	Nelson		13
	4	Dave	25000	12	4	Polk	5000	15
	5	Ed	21000	14	6	Quincy	34000	
					7	Roosevelt	.a	16
DSmerge2								
rowID		lastName	income	famID2	in DSmerge3, _	and .a are miss	sing values that have	been defined in
	1	Lincoln	31000	11	advance by the	user (this table	was created in SAS).
	2	Madison	32000	12				
	3	Nelson	33000	13				
	6	Quincy	34000	14				
DSmerge2a								
caseID		lastName	income	famID2				
	1	Lincoln	31000	11				
	2	Madison	32000	12				
	3	Nelson	33000	13				
	6	Quincy	34000	14				
DSmerge2b								
rowID		lastName	income	famID				
	1	Lincoln	31000	11				
	2	Madison	32000	12				
	3	Nelson	33000	13				
	6	Quincy	34000	14				

caseID	firstName	income	famID	rowID	lastName	famID2
1	Abe	15000	11	1	Lincoln	11
2	Barry		11	2	Madison	12
3	Chuck	17000	12	3	Nelson	13
4	Dave	25000	12	6	Quincy	14
5	Ed	21000	14			
SPSS	MATCH FILES / /FILE='DSmerge	│ FILE=='DSmerge1 e2.sav'	l.sav'			
Stata	use "DSmerge1. merge 1:1 _n us	dta" ing "DSmerge2.dt	a"			
SDTL						
MergeDatasets						
	MergeByVariabl	es				
	FirstVariable					
	LastVariable					
	MergeFileDescri	iption	MergeFileDescri	ption		
	FileName	DSmerge1.sav	FileName	DSmerge2.sav		
	MergeType	Sequential	MergeType	Sequential		
	MergeFlagVarial	ble	MergeFlagVariab	ole		
	RenameVariable	2S	RenameVariable	S		
	update	Master	update	FillNew		
	NewRow	TRUE	NewRow	TRUE		
	KeepVariables		KeepVariables			
	DropVariables		DropVariables			

caseID	firstName	income	famID	rowID	lastName	famID2
1	Abe	31000	11	1	Lincoln	11
2	Barry	32000	11	2	Madison	12
3	Chuck	33000	12	3	Nelson	13
4	Dave	34000	12	6	Quincy	14
5	Ed	21000	14			
SAS	merge DSmerge	1 DSmerge2;				
SDTL						
MergeDatasets						
	MergeByVariabl	es				
	FirstVariable					
	LastVariable					
	MergeFileDescri	ption	MergeFileDescri	ption		
	FileName	DSmerge1.sav	FileName	DSmerge2.sav		
	MergeType	Sequential	MergeType	Sequential		
	MergeFlagVarial	ole	MergeFlagVariab	ole		
	RenameVariable	2S	RenameVariable	S		
	update	Master	update	Replace		
	NewRow	TRUE	NewRow	TRUE		
	KeepVariables		KeepVariables			
	DropVariables		DropVariables			

caseID	firstName	income	famID	rowID	lastName	famID2
1	Abe	15000	11		Lincoln	11
2	Barry	32000	11	2	Madison	12
3	Chuck	17000	12	3	Nelson	13
4	Dave	25000	12	6	Quincy	14
5	Ed	21000	14	NaN	NaN	NaN
Python	DSmerge1.comb	pine_first(DSmerge	e2)			
SDTL						
MergeDatasets						
	MergeByVariabl	es				
	FirstVariable					
	LastVariable					
	MergeFileDescri	iption	MergeFileDescri	ption		
	FileName	DSmerge1.sav	FileName	DSmerge2.sav		
	MergeType	Sequential	MergeType	Sequential		
	MergeFlagVarial	ble	MergeFlagVariab	ole		
	RenameVariable	2S	RenameVariable	S		
	update	Master	update	UpdateMissing		
	NewRow	TRUE	NewRow	TRUE		
	KeepVariables		KeepVariables			
	DropVariables		DropVariables			

Does not repl	ace or update o	verlapping columns	s; instead both a	re included	with no rename							
caseID	firstName	income	famID	rowID	lastName	income	famID2					
	1 Abe	15000	11		1 Lincoln	31000	1	1				
	2 Barry	NaN	11		2 Madison	32000	1	2				
	3 Chuck	17000	12		3 Nelson	33000	1	3				
	4 Dave	25000	12		6 Quincy	34000	1	4				
	5 Ed	21000	14	NaN	NaN	NaN	NaN					
Python	pd.concat([DS	Smerge1, DSmerge2]], axis=1)									
OUT OF S	COPE for S	SDTL										
This Python co	mmand creates	a dataframe with a di	uplicate column n	ame. This	can be prevented by	using the verify_inte	grity paramter.	Since the ou	tput dataframe would t	e unusable, it is no	t in scope for SDT	L.

caseID	firstName	income	famID	lastName	famID2						
1	Abe	15000	11	Lincoln	11						
2	2 Barry		11	Madison	12	2					
3	Chuck	17000	12	Nelson	13	3					
4	Dave	25000	12								
5	Ed	21000	14								
6	3	34000		Quincy	14						
SPSS	/FILE='DSmerge /RENAME rowID /BY caseID	=caseID									
Stata	merge 1:1 casei	d using "DSmerge	2a.dta", nogener	ate		"nogenerate" su	ppresses the crea	tion of a _merge	variable. See Me	rgeFlagVariable 3 and	
SAS	update DSme by caseID;	rge2a DSmerge1	UPDATEMODE=	nomissingcheck;		the contents are the same, but this results in the following column order: caseID, lastName, income, famID2, firstName, famID					
SDTL											
MergeDatasets											
	MergeByVariabl	caseID									
	FirstVariable										
	LastVariable										
	MergeFileDescri	ption	MergeFileDescri	ption							
	FileName	DSmerge1.sav	FileName	DSmerge2.sav							
	MergeType	OneToOne	MergeType	OneToOne							
	MergeFlagVarial	3	MergeFlagVariab								
	RenameVariables Rename\		RenameVariable	oldVariable: "row newVariable: "ca							
	update	Master	update	FillNew							
	NewRow	TRUE	NewRow	TRUE							
	KeepVariables		KeepVariables								
					1						

caseID	firstName	income	famID2	rowID	lastName
1	Abe	15000	11	1	Lincoln
2	Barry		11		
3	Chuck	17000	12	2	Madison
4	Dave	25000	12		
		33000	13	3	Nelson
5	Ed	21000	14	6	Quincy
SPSS	MATCH FILES /F /rename famID=f /FILE='DSmerge: /by famID2		.sav'		
SDTL					
MergeDatasets					
Mergebatasets	MergeByVariable	famID2			
	FirstVariable	lambz			
	LastVariable				
	MergeFileDescrip	otion	MergeFileDescrip	otion	
	FileName	DSmerge1.sav	FileName	DSmerge2.sav	
	MergeType	OneToOne	MergeType	OneToOne	
	MergeFlagVariab		MergeFlagVariab		
		oldVariable: "fam newVariable: "far			
	update	Master	update	FillNew	
	NewRow	TRUE	NewRow	TRUE	
	KeepVariables		KeepVariables		
	DropVariables		DropVariables		

Full join, no By	variable							
caseID	firstName	income	famID	rowID	lastName	famID2		
•	1 Abe	15000	11	NA	NA	NA		
	2 Barry	NA	11	NA	NA	NA		
	Chuck	17000	12	NA	NA	NA		
4	1 Dave	25000	12	NA	NA	NA		
	5 Ed	21000	14	NA	NA	NA		
NA	NA	31000	NA	1	Lincoln	11		
NA	NA	32000	NA	2	Madison	12		
NA	NA	33000	NA	3	Nelson	13		
NA	NA	34000	NA	6	Quincy	14		
Python	DSmerge1.mer	ge(DSmerge2, hov	v="outer")					
R	ds4 <- full_join(DataMerge1, Data	Merge2)					
SDTL								
MergeDatasets								
	MergeByVariab	les						
	FirstVariable							
	LastVariable							
	MergeFileDesc	ription	MergeFileDescri	ption				
	FileName	DSmerge1.sav	FileName	DSmerge2.sav				
	MergeType	Cartesian	MergeType	Cartesian				
	MergeFlagVaria	able	MergeFlagVarial	ole				
	RenameVariabl		RenameVariable	es .				
	update	Master	update	FillNew				
	NewRow	TRUE	NewRow	TRUE				
	KeepVariables		KeepVariables					
	DropVariables		DropVariables					

Full join, no rep	eats in By varial	ole											
caseID	firstName	income.x	famID	lastName	income.y	famID2							
1	Abe	15000	11	Lincoln	31000	11							
2	Barry	NA	11	Madison	32000	12							
3	Chuck	17000	12	Nelson	33000	13							
4	Dave	25000	12	NA	NA	NA							
5	Ed	21000	14	NA	NA	NA							
6	NA	NA	NA	Quincy	34000	14							
Python	DSmerge1.merg	OSmerge1.merge(DSmerge2a, on="caseID", how="outer", suffixes=(".x", ".y"))											
R	ds3 <- full_join(E	DataMerge1, Datal	Merge2a, by=""ca	seID"")									
SDTL													
MergeDatasets													
	MergeByVariable	caseID											
	FirstVariable												
	LastVariable												
	MergeFileDescri	ption	MergeFileDescri	otion									
	FileName	DSmerge1.sav	FileName	DSmerge2.sav									
	MergeType	Cartesian	MergeType	Cartesian									
	MergeFlagVarial	ble	MergeFlagVariab	ole									
	RenameVariable	oldVariable: "inco e newVariable: "inc		oldVariable: "inc newVariable: "in									
	update	Master	update	FillNew									
	NewRow	TRUE	NewRow	TRUE									
	KeepVariables		KeepVariables										
	DropVariables		DropVariables										

Full join, with re	epeats in By var	iable					
caseID	firstName	income.x	famID	rowID	lastName	income.y	
	Abe	15000	11	1	Lincoln	31000	
2	Barry	NA	11		Lincoln	31000	
	Chuck	17000	12	2	Madison	32000	
4	Dave	25000	12	2	Madison	32000	
5	Ed	21000	14	6	Quincy	34000	
NA	NA	NA	13	3	Nelson	33000	
Python	DSmerge1.merg	ge(DSmerge2b, on="famID"	, how="outer", suf	fixes=(".x", ".y"))		sort=True causes the	he rows to be ordered by faml
R	ds13 <- full_join	(DataMerge1, DataMerge2b	o, by="famID")				
Stata	use "DSmerge1 merge m:m fam	.dta" , clear id using "DSmerge2c.dta"					
SDTL							
MergeDatasets							
	MergeByVariab	le caseID					
	FirstVariable						
	LastVariable						
	MergeFileDescr	iption	MergeFileDescri	ption			
	FileName	DSmerge1.sav	FileName	DSmerge2.sav			
	MergeType	Cartesian	MergeType	Cartesian			
	MergeFlagVaria	ble	MergeFlagVarial	ole			
	RenameVariable	oldVariable: "income" newVariable: "income.x"	RenameVariable	oldVariable: "inc newVariable: "inc			
	update	Master	update	FillNew			
	NewRow	TRUE	NewRow	TRUE			
	KeepVariables		KeepVariables				
	DropVariables		DropVariables				

Left join, no rep	eates in By varia	ible					
	caseID	firstName	income.x	famID	lastName	income.y	famID2
1	1	Abe	15000	11	Lincoln	31000	11
2	2	Barry	NA	11	Madison	32000	12
3	3	Chuck	17000	12	Nelson	33000	13
4	4	Dave	25000	12	NA	NA	NA
5	5	Ed	21000	14	NA	NA	NA
Python	DSmerge1.merg	e(DSmerge2a, on	="caseID", how="	left", suffixes=(".x	", ".y"))		
R	ds5 <- left_join(D	ataMerge1, DataN	Merge2a, by="cas	eID")			
SDTL							
MergeDatasets							
	MergeByVariable	caseID					
	FirstVariable						
	LastVariable						
	MergeFileDescrip	otion	MergeFileDescrip	ption			
	FileName	DSmerge1.sav	FileName	DSmerge2.sav			
	MergeType	Cartesian	MergeType	Cartesian			
	MergeFlagVariab	ole	MergeFlagVariab	ole			
	RenameVariable	oldVariable: "inconewVariable: "inco		oldVariable: "inco newVariable: "inc			
	update	Master	update	FillNew			
	NewRow	TRUE	NewRow	FALSE			
	KeepVariables		KeepVariables				
	DropVariables		DropVariables				

Left join, with re	epeates in By var	riable				
caseID	firstName	income.x	famID	rowID	lastName	income.y
1	Abe	15000	11	1	Lincoln	31000
2	Barry	NA	11	1	Lincoln	31000
3	Chuck	17000	12	2	Madison	32000
4	Dave	25000	12	2	Madison	32000
5	Ed	21000	14	6	Quincy	34000
Python	DSmerge1.merg	e(DSmerge2b, on	="famID", how="le	eft", suffixes=(".x",	, ".y"))	
R	ds15 <- left_join(DataMerge1, Data	aMerge2b, by="fa	mID")		
SDTL						
MergeDatasets						
	MergeByVariable	famID				
	FirstVariable					
	LastVariable					
	MergeFileDescri	ption	MergeFileDescrip	otion		
	FileName	DSmerge1.sav	FileName	DSmerge2.sav		
	MergeType	Cartesian	MergeType	Cartesian		
	MergeFlagVariab	ole	MergeFlagVariab	ole		
	RenameVariable	oldVariable: "inco newVariable: "inc	RenameVariable	oldVariable: "inconewVariable: "inco		
	update	Master	update	FillNew		
	NewRow	TRUE	NewRow	FALSE		
	KeepVariables		KeepVariables			
	DropVariables		DropVariables			

J, J	peats in By varia			familia			faIDO
	caseID	firstName	income.x	famID	lastName	income.y	famID2
	1	Abe	15000	11	Lincoln	31000	
	2	Barry	NA	11	Madison	32000	12
	3	Chuck	17000	12	Nelson	33000	13
	6	NA	NA	NA	Quincy	34000	14
Python	DSmerge1.merg	e(DSmerge2a, on	="caseID", how="	right", suffixes=(".	x", ".y"))		
R	ds6 <- right_join(DataMerge1, Data	aMerge2a, by="ca	aseID")			
SDTL							
MergeDatasets							
	MergeByVariable	caseID					
	FirstVariable						
	LastVariable						
	MergeFileDescrip	ption	MergeFileDescri	ption			
	FileName	DSmerge1.sav	FileName	DSmerge2.sav			
	MergeType	Cartesian	MergeType	Cartesian			
	MergeFlagVariab	ole	MergeFlagVariab	ole			
	RenameVariable	oldVariable: "inconewVariable: "inco		oldVariable: "inconewVariable: "inco			
	update	Merge	update	FillNew			
	NewRow	FALSE	NewRow	TRUE			
	KeepVariables		KeepVariables				
	DropVariables		DropVariables				

Right join repea	ats in By variable)								
caseID	firstName	income.x	famID	rowID	lastName	income.y				
1	Abe	15000	11	1	Lincoln	31000				
2	Barry	NA	11	1	Lincoln	31000				
	Chuck	17000	12	2	Madison	32000				
4	Dave	25000	12	2	Madison	32000				
NA	NA	NA	13	3	Nelson	33000				
5	Ed	21000	14	6	Quincy	34000				
Python	DSmerge1.merge(DSmerge2b, on="famID", how="right", sort=True,				uffixes=(".x", ".y"))	if sort=False, th	e row with missing	g caseID would cor	ne last
R	ds16 <- right_join	n(DataMerge1, Da	Merge1, DataMerge2b, by="famID")							
SDTL										
MergeDatasets										
-	MergeByVariable caseID									
	FirstVariable									
	LastVariable									
	MergeFileDescri	ption	MergeFileDescrip	otion						
	FileName		FileName	DSmerge2.sav						
	MergeType	Cartesian	MergeType	Cartesian						
	MergeFlagVarial	ole	MergeFlagVariab	ole						
		oldVariable: "inco	RenameVariable	oldVariable: "inconewVariable: "inco	ome" come.y"					
	update	FillNew	update	Master	-					
	NewRow	FALSE	NewRow	TRUE						
	KeepVariables		KeepVariables							
	DropVariables		DropVariables							

Inner join no rep	peats in By varia	ble				
caseID	firstName	income.x	famID	lastName	income.y	famID2
1	Abe	15000	11	Lincoln	31000	11
2	Barry	NA	11	Madison	32000	12
3	Chuck	17000	12	Nelson	33000	13
Python	DSmerge1.merg	e(DSmerge2a, on	="caseID", suffixe	s=(".x", ".y"))		
R	ds7 <- inner_join					
SDTL						
MergeDatasets						
0.000	MergeByVariable	caseID				
	FirstVariable					
	LastVariable					
	MergeFileDescrip	otion	MergeFileDescrip	otion		
	FileName	DSmerge1.sav	FileName	DSmerge2.sav		
	MergeType	OneToOne	MergeType	OneToOne		
	MergeFlagVariab	le	MergeFlagVariab	le		
	RenameVariable	oldVariable: "inconewVariable: "inco	RenameVariable	oldVariable: "inconewVariable: "inco		
	update	Master	update	FillNew		
	NewRow	FALSE	NewRow	FALSE		
	KeepVariables		KeepVariables			
	DropVariables		DropVariables			

Inner join repeat	s in By variable					
caseID	firstName	income.x	famID	rowID	lastName	income.y
1	Abe	15000	11	1	Lincoln	31000
2	Barry	NA	11	1	Lincoln	31000
3	Chuck	17000	12	2	Madison	32000
4	Dave	25000	12	2	Madison	32000
5	Ed	21000	14	6	Quincy	34000
Python	DSmerge1.merg	e(DSmerge2b, on	="famID", suffixes	s=(".x", ".y"))		
R	ds17 <- inner_joi	in(DataMerge1, D	ataMerge2b, by="	famID")		
SDTL						
MergeDatasets						
	MergeByVariable	caseID				
	FirstVariable					
	LastVariable					
	MergeFileDescri	ption	MergeFileDescri	ption		
	FileName	DSmerge1.sav	FileName	DSmerge2.sav		
	MergeType	OneToOne	MergeType	OneToOne		
	MergeFlagVarial	ole	MergeFlagVariat	ole		
	RenameVariable	oldVariable: "inconewVariable: "inco		oldVariable: "inconewVariable: "inco		
	update	Master	update	FillNew		
	NewRow	FALSE	NewRow	FALSE		
	KeepVariables		KeepVariables			
	DropVariables		DropVariables			

semi join, no re	peats in By varia	ble				
caseID	firstName	income	famID			
1	Abe	15000	11			
2	Barry	NA	11			
3	Chuck	17000	12			
R	ds8 <- semi_join((DataMerge1, Dat	aMerge2a, by="c	aseID")		
SDTL						
MergeDatasets						
	MergeByVariable	caseID				
	FirstVariable					
	LastVariable					
	MergeFileDescrip	otion	MergeFileDescription			
	FileName	DSmerge1.sav	FileName	DSmerge2.sav		
	MergeType	Cartesian	MergeType	Cartesian		
	MergeFlagVariab	le	MergeFlagVariat	ole		
	RenameVariables	S	RenameVariable	S		
	update	Master	update	Ignore		
	NewRow	FALSE	NewRow	FALSE		
	KeepVariables		KeepVariables			
	DropVariables		DropVariables			

semi join, with r	epeats in By variab	le			
caseID	firstName	income	famID		
1	Abe	15000	11		
2	Barry	NA	11		
3	Chuck	17000	12		
4	Dave	25000	12		
5	Ed	21000	14		
R	ds18 <- semi_join(D	ataMerge1, DataN	/lerge2b, by="fam	ID")	
SDTL					
MergeDatasets					
	MergeByVariables	caseID			
	FirstVariable				
	LastVariable				
	MergeFileDescription	n	MergeFileDescription		
	FileName	DSmerge1.sav	FileName	DSmerge2.sav	
	MergeType	Cartesian	MergeType	Cartesian	
	MergeFlagVariable		MergeFlagVariab	ole	
	RenameVariables		RenameVariable	S	
	update	Master	update	FillNew	
	NewRow	FALSE	NewRow	FALSE	
	KeepVariables		KeepVariables		
	DropVariables		DropVariables		

caseID	firstName	income	famID			
4	Dave	25000	12			
5	Ed	21000	14			
R	ds9 <- anti_join([DataMerge1, Data	Merge2a, by="ca	seID")		
SDTL						
MergeDatasets						
	MergeByVariable	caseID				
	FirstVariable					
	LastVariable					
	MergeFileDescrip	otion	MergeFileDescription			
	FileName	DSmerge1.sav	FileName	DSmerge2.sav		
	MergeType	Unmatched	MergeType	Unmatched		
	MergeFlagVariab	ole	MergeFlagVariab	ole		
	RenameVariable	S	RenameVariable	S		
	update	Master	update	Ignore		
	NewRow	TRUE	NewRow	FALSE		
	KeepVariables		KeepVariables			
	DropVariables		DropVariables			

caseid	firstn~e	income	famid2	lastname	_merge
1	Abe	15000	11	Lincoln	matched
2	Barry		11	Lincoln	matched
3	Chuck	17000	12	Madison	matched
4	Dave	25000	12	Madison	matched
5	Ed	21000	14	Quincy	matched
3		33000	13	Nelson	using only
Stata	merge m:1 famic	using "DSmerge:	2a.dta"		
SDTL					
MergeDatasets					
	MergeByVariable	famID			
	FirstVariable				
	LastVariable				
	MergeFileDescri	ption	MergeFileDescrip		
	FileName	DSmerge1.sav	FileName	DSmerge2.sav	
	MergeType	ManyToOne	MergeType	OneToMany	
	MergeFlagVarial	ole	MergeFlagVariab	ole	
	RenameVariable	S	RenameVariable	oldVariable: "fan newVariable: "fa	
	update	Master	update	FillNew	
	NewRow	TRUE	NewRow	TRUE	
	KeepVariables		KeepVariables		
	DropVariables		DropVariables		

caseid	firstname	income	famid	lastname	famid2			
	Abe	15000	11	Lincoln	11			
2	Barry	32000	11	Madison	12			
;	Chuck	17000	12	Nelson	13			
4	Dave	25000	12					
	Ed	21000	14					
(3	34000		Quincy	14			
Stata	merge 1:1 case	id using "DSmerge	2a.dta" , update					
SAS	update DSmerge2a DSmerge1 UPDATEMODE=missingcheck; by caseID;				contents are the same, but this results in the following column caseID, lastName, income, famID2, firstName, famID			
SDTL								
MergeDatasets								
	MergeByVariab	l∈ caseID						
	FirstVariable							
	LastVariable							
	MergeFileDescr	ription	MergeFileDescri	ption				
	FileName	DSmerge1.sav	FileName	DSmerge2.sav				
	MergeType	OneToOne	MergeType	OneToOne				
	MergeFlagVaria	ible	MergeFlagVariab	ole				
	RenameVariables RenameVariables							
	update	UpdateNulls	update	UpdateMissing				
	NewRow	TRUE	NewRow	TRUE				
	KeepVariables		KeepVariables					
	DropVariables		DropVariables					

caseid	firstname	income	famid	lastname	famid2			
1	Abe	31000	11	Lincoln	11			
2	Barry	32000	11	Madison	12			
3	Chuck	33000	12	Nelson	13			
4	Dave	25000	12					
5	Ed	21000	14					
6		34000		Quincy	14			
Stata	merge 1:1 caseid using "DSmerg		2a.dta" , update r	eplace				
SAS	merge DSmer by caseID;	ge1 DSmerge2a;				mergeType sl	hould probably be	e "SASmatchMerge
SDTL								
MergeDatasets								
	MergeByVariab	l∈ caseID						
	FirstVariable							
	LastVariable							
	MergeFileDescr	iption	MergeFileDescri	ption				
	FileName	DSmerge1.sav	FileName	DSmerge2.sav				
	MergeType	OneToOne	MergeType	OneToOne				
	MergeFlagVaria	ble	MergeFlagVarial	ole				
	RenameVariables		RenameVariables					
	update	Master	update	Replace				
	NewRow	TRUE	NewRow	TRUE				
	KeepVariables		KeepVariables					
	DropVariables		DropVariables					

Ignore columns	not in the mast	er dataframe, and	l update values t	hat are missing	in the master d	ataframe for Up	dateping columns
caseID	firstName	income	famID				
1	Abe	15000	11				
2	Barry	32000	11				
3	Chuck	17000	12				
4	Dave	25000	12				
5	Ed	21000	14				
Python	DSmerge1.upda	te(DSmerge2, ove	erwrite=False)	this acts in-place	on DSmerge1		
SDTL							
MergeDatasets							
	MergeByVariabl	es					
	FirstVariable						
	LastVariable						
	MergeFileDescri	ption	MergeFileDescri	ption			
	FileName	DSmerge1.sav	FileName	DSmerge2.sav			
	MergeType	Sequential	MergeType	Sequential			
	MergeFlagVarial	ole	MergeFlagVariat	ole			
	RenameVariable	2S	RenameVariable	S			
	update	Master	update	UpdateMissing			
	NewRow	TRUE	NewRow	FALSE			
	KeepVariables		KeepVariables				
	DropVariables		DropVariables	lastName, famID	2, rowID		

caseID	firstName	income	famID	lastName				
1	Abe	_		Lincoln				
2	Barry	32000	Α	Madison				
3	Chuck		13	Nelson				
4	Dave	5000	15	Polk				
5	Ed	21000	14					
6	3	34000		Quincy				
7	•	Α	16	Roosevelt				
SAS	by caseID;	ge i Domergeo o	PDATEMODE=no	ornissingeneek,				
SDTI	MergeDatasets							
SDTL	MergeDatasets MergeByVariable	VariableSymbol	Evnression					
SDTL	_	VariableSymbol						
SDTL	MergeByVariable	variableName	caseID	MergeFileDesc	rintion			
SDTL	_	variableName MergeFileDescr	caseID iption	MergeFileDeso				
SDTL	MergeByVariable	variableName MergeFileDescr fileName	caseID iption DSmerge1	fileName	DSmerge3			
SDTL	MergeByVariable	variableName MergeFileDescr fileName mergeType	caseID iption DSmerge1 OneToOne	fileName mergeType	DSmerge3 OneToOne			
SDTL	MergeByVariable	variableName MergeFileDescr fileName mergeType update	caseID iption DSmerge1 OneToOne FillNew	fileName mergeType update	DSmerge3 OneToOne Master			
SDTL	MergeByVariable	variableName MergeFileDescr fileName mergeType update newRow	caseID iption DSmerge1 OneToOne	fileName mergeType	DSmerge3 OneToOne			

With missingch	eck option, non-	missing values i	n the SASUpdate	e dataframe wil	l replace values	in master, as will	user-defined m	issing, but system	missing will not
caseID	firstName	income	famID	lastName					
1	Abe		11	Lincoln					
2	Barry	32000	Α	Madison					
3	Chuck	17000	13	Nelson					
4	Dave	5000	15	Polk					
5	Ed	21000	14						
6		34000		Quincy					
7		Α	16	Roosevelt					
SAS	data DSmerge4; update DSmer by caseID; MergeDatasets		JPDATEMODE=m	nissingcheck;					
SDIL	mergeByVariable	VariableSymbo	 Evnrossion						
	mergeby variable	variableName	caseID						
	MergeDatasets	MergeFileDesci		MergeFileDes	cription				
		fileName	DSmerge1	fileName	DSmerge3				
		mergeType	OneToOne	mergeType	OneToOne				
		update	UpdateMissing	Update	Master				
		newRow	TRUE	newRow	TRUE				
	consumesDatafra	DSmerge1	Dsmerge3						
	producesDatafra	DSmerge4							

caseID	lastname	income	famID2	firstname	famID
1	Lincoln	15000	11	Abe	11
2	Madison		12	Barry	11
3	Nelson	17000	13	Chuck	12
4		25000		Dave	12
5		21000		Ed	14
6	Quincy	34000	14		
SAS	update DSmerge by caseID;	e2a DSmerge1 UF	PDATEMODE=no	missingcheck;	
SDTL	MergeDatasets				
	mergeByVariables	VariableSymbol	Expression		
		variableName	caseID		
	MergeDatasets	MergeFileDescr	iption	MergeFileDescr	ription
		fileName	DSmerge2a	fileName	DSmerge1
		mergeType	OneToOne	mergeType	OneToOne
		update	FillNew	update	Master
		newRow	TRUE	newRow	TRUE

caseID	lastname	income	famID2	firstname	famID
1	Lincoln	15000	11	Abe	11
2	Madison	32000	12	Barry	11
3	Nelson	17000	13	Chuck	12
4		25000		Dave	12
5		21000		Ed	14
6	Quincy	34000	14		
SAS	update DSmer by caseID;	ge2a DSmerge1	UPDATEMODE=r	missingcheck;	
SDTL	MergeDatasets				
	mergeByVariable	VariableSymbol	Expression		
		variableName	caseID		
	MergeDatasets	MergeFileDescr	iption	MergeFileDesci	ription
		fileName	DSmerge2a	fileName	DSmerge1
		mergeType	Cartesian	mergeType	Cartesian
		update	UpdateMissing	update	Master
		newRow	TRUE	newRow	TRUE

caseID	lastname	income	famID2					
1	Lincoln	15000	11					
2	Madison	32000	12					
3	Nelson	17000	13					
6	Quincy	34000	14					
SAS	data DSmerge2a modify DSmer by caseID;	u; ge2a DSmerge1	UPDATEMODE=I	missingcheck;	this acts on DSn	nerge2a in-place		
SDTL	MergeDatasets							
	mergeByVariable	VariableSymbo	Expression					
		variableName	caseID					
	MergeDatasets	MergeFileDescr	ription	MergeFileDesc	ription			
		fileName	DSmerge2a	fileName	DSmerge1			
		mergeType	Cartesian	mergeType	Cartesian			
		update	UpdateMissing	update	Master			
		newRow	TRUE	newRow	FALSE			
				dropVariables	VariableSymbo	lExpression	VariableSymbo	lExpression
					variableName	firstName	variableName	famID
	consumesDatafra	DSmerge2a	DSmerge1					
	producesDatafra	DSmerge2a						

caseID	lastname	income	famID2					
1	Lincoln	15000	11					
2	Madison		12					
3	Nelson	17000	13					
6	Quincy	34000	14					
SAS	data DSmerge2a; modify DSmerge2a by caseID;	DSmerge1 UPDA	ATEMODE=nomis	ssingcheck;	this acts on DSn	nerge2a in-place		
SDTL	MergeDatasets							
	mergeByVariables	VariableSymbo	Expression					
		variableName	caseID					
	MergeDatasets	MergeFileDesci	ription	MergeFileDesc	ription			
		fileName	DSmerge2a	fileName	DSmerge1			
		mergeType	OneToOne	mergeType	OneToOne			
		update	FillNew	update	Master			
		newRow	TRUE	newRow	FALSE			
				dropVariables	VariableSymbo	IExpression	VariableSymbo	IExpression
					variableName	firstName	variableName	famID
	consumesDataframe	DSmerge2a	DSmerge1				_	
	producesDataframe	DSmerge2a						

caseID	firstName	income	famID		
1	Abe	31000	11		
2	Barry	32000	11		
3	Chuck	33000	12		
4	Dave	34000	12		
5	Ed	21000	14		
SAS	data DSmerge1; modify DSmer by caseID;	rge1 DSmerge2a ;			this acts on DSmerge1 in-place
SDTL					
MergeDatasets					
	MergeByVariabl	caseID			
	FirstVariable				
	LastVariable				
	MergeFileDescri	iption	MergeFileDescri	ption	
	FileName	DSmerge1.sav	FileName	DSmerge2a.sav	
	MergeType	OneToOne	MergeType	OneToOne	
	MergeFlagVarial	ble	MergeFlagVarial	ble	
	RenameVariable	es ·	RenameVariable	es ·	
	update	Master	update	Replace	
	NewRow	TRUE	NewRow	FALSE	
	KeepVariables		KeepVariables		
	DropVariables		DropVariables	lastName, famID	2

caseID	firstName	income	famID			
1	Abe	31000	11			
2	Barry	32000	11			
3	Chuck	33000	12			
4	Dave	34000	12			
5	Ed	21000	14			
Python	DSmerge1.upda	ate(DSmerge2)			this acts on DSm	erge1 in-place
SDTL						
MergeDatasets						
	MergeByVariab	les				
	FirstVariable					
	LastVariable					
	MergeFileDesci	ription	MergeFileDescri	ption		
	FileName	DSmerge1.sav	FileName	DSmerge2.sav		
	MergeType	Sequential	MergeType	Sequential		
	MergeFlagVaria	ible	MergeFlagVariab	ole		
	RenameVariabl	es	RenameVariable	S		
	update	Master	update	Replace		
	NewRow	TRUE	NewRow	FALSE		
	KeepVariables		KeepVariables			
	DropVariables		DropVariables	lastName, famID	2, rowID	

caseID	firstName	income	famID	rowID	lastName
1	7 1.0 0	15000	11		Lincoln
	Barry	4-000	11		Lincoln
3		17000	12		Madison
4		25000	12		Madison
5	Ed	21000	14	6	Quincy
SPSS	MATCH FILES /F /TABLE='DSmerous /RENAME famID /BY famID	FILE=='DSmerge´ ge2.sav' 2=famID	l.sav'		
SDTL					
MergeDatasets					
	MergeByVariable	famID			
	FirstVariable				
	LastVariable				
	MergeFileDescrip	otion	MergeFileDescrip	otion	
	FileName	DSmerge1.sav	FileName	DSmerge2.sav	
	MergeType	ManyToOne	MergeType	OneToMany	
	MergeFlagVariab	le	MergeFlagVariab	ole	
	RenameVariables	S	RenameVariable	oldVariable: "fam newVariable: "fai	
	update	Master	update	Ignore	
	NewRow	TRUE	NewRow	FALSE	
	KeepVariables		KeepVariables		
	DropVariables		DropVariables		

caseID	firstName	income	famID	lastName	famID2	inDS1	inDS2
1	Abe	15000	11	Lincoln	11	1	
2	Barry		11	Madison	12	1	
3	Chuck	17000	12	Nelson	13	1	
4	Dave	25000	12			1	
5	Ed	21000	14			1	
6		34000		Quincy	14	0	
SPSS	MATCH FILES /I /IN inDS1 /FILE='DSmerge /RENAME rowID /IN inDS2 /BY caseID		.sav'				
SDTL							
MergeDatasets							
	MergeByVariable	caseID					
	FirstVariable						
	LastVariable						
	MergeFileDescri	ption	MergeFileDescrip	otion			
	FileName	DSmerge1.sav	FileName	DSmerge2.sav			
	MergeType	OneToOne	MergeType	OneToOne			
	MergeFlagVariab	inDS1	MergeFlagVariab	inDS2			
	RenameVariable	S	RenameVariable	oldVariable: "rov newVariable: "c			
	update	Master	update	FillNew			
	NewRow	TRUE	NewRow	TRUE			
	KeepVariables		KeepVariables				

Dron\/ari	ables Dror	nVariables	
Diopvari	anics) variables	

caseID	firstName	income_x	famID	lastName	income_y	famID2	_merge
1	Abe	15000	11	Lincoln	31000	11	both
2	Barry	NaN	11	Madison	32000	12	both
3	Chuck	17000	12	Nelson	33000	13	both
4	Dave	25000	12	NaN	NaN	NaN	left_only
5	Ed	21000	14	NaN	NaN	NaN	left_only
6	NaN	NaN	NaN	Quincy	34000	14	right_only
Python	DSmerge1.merg	∣ e(DSmerge2.rena □	me(columns={"ro	wID": "caseID"}),	on="caseID", how	="outer", indicator	=True)
MergeDatasets							
	MergeByVariable	caseID					
	FirstVariable						
	LastVariable						
	MergeFileDescri	ption	MergeFileDescription				
	FileName	DSmerge1.sav	FileName	DSmerge2.sav			
	MergeType	OneToOne	MergeType	OneToOne			
	MergeFlagVariab	_merge	MergeFlagVariab	ole			
	RenameVariable	S	RenameVariable	oldVariable: "rov newVariable: "ca			
	update	Master	update	FillNew			
	NewRow	TRUE	NewRow	TRUE			
	KeepVariables		KeepVariables				
	DropVariables		DropVariables				

caseID	firstName	income	famID	lastName	famID2							
1	Abe	15000	11	Lincoln	11							
2	Barry		11	Madison	12							
3	Chuck	17000	12	Nelson	13							
4	Dave	25000	12									
5	Ed	21000	14									
6	3	34000		Quincy	14							
Stata	merge 1:1 case	id using "DSmerge	2a.dta"			Stata creates a	rariable named _	merge with the fo	ollowing codes			
						Use SDTL SetVa	alueLabels to as	sign labels to _m	erge			
DTL							value	label	observation app	eared in master fi	le only	
MergeDatasets						1		1 master	observation app	eared in other file	only	
	MergeByVariab	l∈ caseID				2	2	2 using	observation app	eared in both files	3	
	FirstVariable					3	;	3 match	observation app	eared in both, mis	ssing values upda	ted
	LastVariable					4	4	match_update	observation app	eared in both, cor	nflicting nonmissir	ıg values
	MergeFileDesci	iption	MergeFileDescri	ption		5	!	match_conflict				
	FileName	DSmerge1.sav	FileName	DSmerge2.sav								
	MergeType	OneToOne	MergeType	OneToOne		See also MergeF	lagVariable 4					
	MergeFlagVaria	b_merge	MergeFlagVariab									
	RenameVariabl	es		oldVariable: "row newVariable: "cas								
	update	Master	update	FillNew								
	NewRow	TRUE	NewRow	TRUE								
	KeepVariables		KeepVariables									
	DropVariables		DropVariables									

caseID	firstName	income	famID	lastName	famID2							
•	1 Abe	15000	11	Lincoln	11							
2	2 Barry		11	Madison	12							
	3 Chuck	17000	12	Nelson	13							
4	1 Dave	25000	12									
į.	5 Ed	21000	14									
(3	34000		Quincy	14							
Stata	merge 1:1 case	id using "DSmerge	2a.dta" , generate	e(Merge_code)		Stata generates	a new variable n	amed Merge_cod	e with the followin	g codes		
						Use SDTL SetVa	lueLabels to as:	sign labels to Merg	e_code			
SDTL							value	label				
MergeDatasets						1		1 master	observation appe	eared in master file	e only	
	MergeByVariab	le caseID				2		2 using	observation appe	eared in other file	only	
	FirstVariable					3		3 match	observation appe	eared in both files		
	LastVariable					4		4 match_update	observation appe	eared in both, miss	sing values upda	ited
	MergeFileDesc	ription	MergeFileDescri	ption		5		5 match_conflict	observation appe	eared in both, con	flicting nonmissi	ng values
	FileName	DSmerge1.sav	FileName	DSmerge2.sav								
	MergeType	OneToOne	MergeType	OneToOne		See also MergeF	lagVariable 3					
	MergeFlagVaria	b Merge_code	MergeFlagVarial	b								
	RenameVariabl	es	RenameVariable	oldVariable: "rowl newVariable: "cas								
	update	Master	update	FillNew								
	NewRow	TRUE	NewRow	TRUE								
	KeepVariables		KeepVariables									
	DropVariables		DropVariables									

caseID	firstName	income	famID	rowID	lastName	firstINfam	lastINfam
1	Abe	15000	11	1	Lincoln	1	
2	Barry		11	1	Lincoln	0	
3	Chuck	17000	12	2	Madison	1	
4	Dave	25000	12	2	Madison	0	
5	Ed	21000	14	6	Quincy	1	
SPSS	MATCH FILES /F /TABLE='DSmern /RENAME famID /first firstINfam /last lastINfam /BY famID .	FILE=='DSmerge´ ge2.sav' i2=famID	I.sav'				
SDTL							
MergeDatasets							
	MergeByVariable	famID					
	FirstVariable	firstINfam					
	LastVariable	lastINfam					
	MergeFileDescri	ption	MergeFileDescrip	otion			
	FileName	DSmerge1.sav	FileName	DSmerge2.sav			
	MergeType	ManyToOne	MergeType	OneToMany			
	MergeFlagVariab		MergeFlagVariab				
	RenameVariable		RenameVariable	oldVariable: "fan newVariable: "fa			
	update	Master	update	Ignore			
	NewRow	TRUE	NewRow	FALSE			
	KeepVariables		KeepVariables				
	DropVariables		DropVariables				

Inner join no re	epeats in by variable								Inner join, repe	ate in by variable									
			famID	rowID	lastName		famID2				income x	famID	rowID	lastName		famID2			
						income_y													
	Abe	15000			1 Lincoln	310		11		Abe	1500			1 Lincoln	31000		11		
	Barry	NaN	11		2 Madison	320		12		Barry	NaN	1		1 Lincoln	31000		11		
3	Chuck	17000	12	! 3	3 Nelson	330	000	13	3	Chuck	1700	0 1:	2 2	2 Madison	32000	1	12		
									4	Dave	2500	0 1:	2 2	2 Madison	32000	1	12		
									5	Ed	2100			6 Quincy	34000		14		
Python	DSmerge1.merge(DS	Smerne? left on:	"caseID" right on=	"rowID")	1				Python	DSmerne1 mern	e(DSmerne2 lef	t on="femID" riol	nt_on="famID2")	1					
SDTL	Domerge r.merge(De	omergez, ren_on	- cascib , rigita_cir	TOWNE)					SDTL	Domerge ranery	(C(DOMCIGGE, ICI	L_OII - IdilliD , rigi	it_on= idinib2)	-					
MergeDatasets									MergeDatasets										
	MergeByVariables	caseID								MergeByVariabl	€ famID								
	FirstVariable									FirstVariable									
	LastVariable									LastVariable									
	MergeFileDescription	1	MergeFileDescripti	ion						MergeFileDescri	intion	MergeFileDesc	ription	1					
	FileName		FileName	DSmerge2.sav						FileName	DSmerge1.sav	FileName	DSmerge2.sav						
														+					
	MergeType	Cartesian	MergeType	Cartesian						MergeType	Cartesian	MergeType	Cartesian	_					
	MergeFlagVariable		MergeFlagVariable	2						MergeFlagVarial		MergeFlagVaria	ible						
		oldVariable: *inci		oldVariable: "inc newVariable: "in	come"					RenameVariable	oldVariable: "in	cd	oldVariable: "inc e newVariable: "in	come"					
	RenameVariables	newVariable: "in	RenameVariables		ncome.y"						newVariable: "ir			ncome.y"					
	update	Master	update	FillNew						update	Master	update	FillNew						
	NewRow	FALSE	NewRow	FALSE						NewRow	FALSE	NewRow	FALSE						
	KeepVariables		KeepVariables							KeepVariables		KeepVariables							
	DropVariables		DropVariables							DropVariables		DropVariables	1	1					
	MergeByNames			roudD.						MergeByNames			famID2	1					
	мегдевуматеѕ		MergeByNames	rowiD						MergebyNames		MergeByName	s ramiD2	_					
	L							-										+	
	peats in by variable								Left join, repeat										
	firstName		famID	rowID	lastName	income_y	famID2		caseID		income_x	famID	rowID	lastName		famID2			
1	Abe	15000	11	1	1 Lincoln	310	000	11	1	Abe	1500	0 1	1 1	1 Lincoln	31000		11		
	Barry	NaN	11		2 Madison		000	12		Barry	NaN	1		1 Lincoln	31000		11		
	Chuck	17000			3 Nelson	330		13		Chuck	1700			2 Madison	32000		12		
	Dave	25000		NaN	NaN	NaN	NaN			Dave	2500			2 Madison	32000		12		
																	12		
5	Ed	21000	14	NaN	NaN	NaN	NaN		5	Ed	2100	0 1-	4 6	6 Quincy	34000		14		
						_													
	DSmerge1.merge(DS	Smerge2, left_on=	="caseID", right_on=	rowID*, how=*le	eft")					DSmerge1.merg	e(DSmerge2, lef	t_on="famID", rigi	ht_on="famID2", h	now="left")					
SDTL									SDTL										
MergeDatasets									MergeDatasets										
_	MergeByVariables	caseID							_	MergeByVariabl	e famID								
	FirstVariable	GUSCID								FirstVariable	- IdilliD								
	LastVariable									LastVariable									
	MergeFileDescription	1	MergeFileDescripti							MergeFileDescri	iption	MergeFileDesc							
	FileName	DSmerge1.sav	FileName	DSmerge2.sav						FileName	DSmerge1.sav	FileName	DSmerge2.sav						
	MergeType	Cartesian	MergeType	Cartesian						MergeType	Cartesian	MergeType	Cartesian						
	MergeFlagVariable		MergeFlagVariable							MergeFlagVarial		MergeFlagVaria							
	recepting variable	oldVariable: *inci	TVICEBELLION VOLIDOR	oldVariable: "inc	nomo"					TVICTBET INB VUTTO	oldVariable: "in	or	oldVariable: *inc	como"					
	RenameVariables	newVariable: "in	RenameVariables	newVariable: "ir	ncome.v"					RenameVariable	newVariable: "in	nd RenameVariabl	e newVariable: "in	ncome.v"					
	update	Master	update	FillNew	1 '					update	Master	update	FillNew						
	NewRow	TRUE	NewRow							NewRow	TRUE	NewRow		+					
		TRUE		FALSE							TRUE		FALSE	-					
	KeepVariables		KeepVariables							KeepVariables		KeepVariables							
	DropVariables		DropVariables							DropVariables		DropVariables							
	MergeByNames		MergeByNames	rowID						MergeByNames		MergeByName	s famID2						
	epeats in by variable								Right join, repe										
caseID	firstName		famID	rowID	lastName	income_y	famID2		caseID	firstName	income_x	famID	rowID	lastName	income_y	famID2			
1	Abe	15000	11	1	1 Lincoln	310	000	11	1	Abe	1500	0 1	1 1	1 Lincoln	31000	1	11		
	Barry	NaN	11		2 Madison	320		12		Barry	NaN	1		1 Lincoln	31000		11		
3	Chuck	17000	12	! 3	3 Nelson	330	000	13	3	Chuck	1700	0 1:	2 2	2 Madison	32000		12		
	NaN		NaN		6 Quincy	340		14		Dave	2500			2 Madison	32000		12		
				`		-				Ed	2100			6 Quincy	34000		14		
									NaN	NaN	NaN	NaN		3 Nelson	33000		13		
									INGIN	· valv	- valv	NOIN	- 3	- Maionii	33000				
D. III.			E IDII		1.00				2 "						_				
Python	DSmerge1.merge(DS	smerge2, left_on=	caseID", right_on=	rowID*, how=*rig	gnt")	1			Python	USmerge1.merg	e(DSmerge2, lef	t_on="tamID", rigi	nt_on="famID2", h	now="right")					
SDTL									SDTL										
MergeDatasets									MergeDatasets										
	MergeByVariables	caseID								MergeByVariabl	€ famID								
	FirstVariable									FirstVariable									
	LastVariable				1					LastVariable		1		1					
	ManagrilaDana		Mana-Fila Dan 111		+					ManagellaD		Manacitac		+					
	MergeFileDescription	no .	MergeFileDescripti		-					MergeFileDescri	puon	MergeFileDesc		-					
	FileName	DSmerge1.sav	FileName	DSmerge2.sav						FileName	DSmerge1.sav	FileName	DSmerge2.sav						
	MergeType		MergeType	Cartesian						MergeType	Cartesian	MergeType	Cartesian						
	MergeFlagVariable		MergeFlagVariable	2						MergeFlagVarial		MergeFlagVaria	ible						
		oldVariable: *inci		oldVariable: "inc	come"						oldVariable: "in	CO	oldVariable: "inc	come"					
	RenameVariables	newVariable: "in	RenameVariables	newVariable: *ir	ncome.y"					RenameVariable	newVariable: "ir	nd RenameVariabl	e newVariable: "in	ncome.y"					
	update	Master	update	FillNew						update	Master	update	FillNew						
	NewRow	FALSE	NewRow	TRUE						NewRow	FALSE	NewRow	TRUE						
	KeepVariables		KeepVariables	oL	1					KeepVariables	TALOL	KeepVariables	- NOL	1					
					-									-					
	DropVariables		DropVariables		-					DropVariables		DropVariables							
	MergeByNames		MergeByNames	rowID						MergeByNames		MergeByName	s famID2						
Outer join, no r	epeats in by variable								Outer join, repe	ats in by variable	e								
	firstName	income x	famID	rowID	lastName	income_y	famID2			firstName	income x	famID	rowID	lastName	income_y	famID2			
	Abe	15000			1 Lincoln	310		11		Abe	1500			1 Lincoln	31000		11		
	7100		11		2 Madison		000	12		Barry	NaN	1		1 Lincoln	31000		11		
1	Parni							14	2	Daily	INGIN	1	1	LITICOTT	31000		11		
1 2	Barry	NaN 47000						42	-	Charale	4900			O Martin	00		10		
1 2 3	Chuck	17000	12	! 3	3 Nelson	330	000	13		Chuck	1700			2 Madison	32000		12		
1 2 3 4	Chuck Dave	17000 25000	12 12	! NaN	3 Nelson NaN	NaN	NaN	13	4	Dave	2500	0 1:	2 2	2 Madison	32000		12		
1 2 3 4	Chuck	17000	12 12	! 3	3 Nelson	330	000	13	4			0 1:	2 2						

NaN	NaN	NaN	NaN	6	Quincy 34000	14	1 1	NaN	NaN NaN	NaN	3 Nelson 33000	13			
Python	DSmerge1.merge(DS	Smerge2, left_on=	="caseID", right_on=	"rowID", how="out	ter")			Python	DSmerge1.merge(DSmerge2, left	on="famID", right	_on="famID2", how="outer")				
SDTL								SDTL							
MergeDatasets							- '	MergeDatasets							
	MergeByVariables	caseID							MergeByVariable famID						
	FirstVariable LastVariable								FirstVariable LastVariable						
			MergeFileDescription						MorgoFiloDoscription	MargaEilaDarasi	stion				
	FileName	DSmerge1 say	MergeFileDescription FileName MergeType MergeFlagVariable	DSmerne2 say					FileName DSmerge1 say	FileName	DSmerge2 say				
	FileName MergeType	Cartesian	MergeType	DSmerge2.sav Cartesian					MergeType Cartesian	MergeType	Cartesian				
	MergeFlagVariable		MergeFlagVariable	1					MergeFlagVariable	MergeFlagVariat	le				
	RenameVariables update NewRow	oldVariable: *inci		oldVariable: "inco	ome"				LastVariable MergeFileDescription FileName DSmerge1, sav MergeType Cartesian MergeTalgvariable newVariable: "in: update Master NewRow TRUE KeepVariables		oldVariable: "income"				
	RenameVariables	newVariable: "in	RenameVariables	newVariable: *inc	come.y"				RenameVariable newVariable: "in	RenameVariable	newVariable: "income.y"				
	NowPow	Master	update NowPow	TOUT					NowPow TDUE	update NowPow	FIINEW				
	KeepVariables	TRUE	KeepVariables	IRUE					KeenVariables	KeepVariables	IRUE				
	DropVariables		DropVariables							DropVariables					
	MergeByNames		MergeByNames	rowID					MergeByNames MergeByNames	MergeByNames	famID2				
			,							. 0. ,					
							-								
							_								
							\perp								
							\perp								
							-								

			-1						
			-						
			-						
			-						
			-						
			-						
			_						
			_						
			-						
			-						
			-1						
			-1						
			-						
			-						
			-						
			-						
			-						

			_						
			_						
			_						
			_						
			_						
			-						
			-						
			-						
			-						
			_						
			-1						
			-						
			-1						
			_						

			1						
			1						
			-						
			-						
			-						
			-						
			-						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			-						
			-						
			_						
			_						

			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						

			1						
			1						
			-						
			-						
			-						
			-						
			-						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			-						
			-						
			_						
			_						

			-						
			-						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			-						
			-						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			1						
			-						

				-							
				-							
				-							
				-							
				1							
				1							
				1							
				1							
				1							
				1							
				-							
				-							
				1							
				1							
				-							
				-							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							

			_						
			_						
			-1-						
			_						
			-1						
			\perp						
			-1						
			-1-						
			_						

Inner Join							Right Join										
	F		r	Land Land		r100		C		fID				r			
caseID			famID rowID			famID2	caseID	firstName				lastName		famID2			
1	1 Abe	15000	11	1 Lincoln	31000	11		Abe	15000			Lincoln	31000				
							NaN	NaN	NaN	NaN	2	Madison	32000	1:			
							NaN	NaN	NaN	NaN	3	Nelson	33000	1:			
							NaN			NaN		Quincy	34000	1-			
							T-Car-t	reare	Huit	reare		Guntoy	04000				
Python	DSmerge1.merg	ge(DSmerge2, left	_on=["caseID", "famID"], right_on=	["rowID", "tamID2"])			Python	DSmerge1.merg	e(DSmerge2, left	_on=["caseID", "far	mID"], right_on=[rowID", "tamID2"], how="right")				
SDTL							SDTL										
MergeDatasets							MergeDatasets										
	MergeByVariabl	caseID famID						MergeByVariabl	caseID famID								
	FirstVariable							FirstVariable									
	LastVariable							LastVariable									
	MergeFileDescri	iption	MergeFileDescription					MergeFileDescri	ption	MergeFileDescrip	ption						
	FileName	DSmerge1.sav	FileName DSmerge2.sav					FileName	DSmerge1.sav	FileName	DSmerge2.sav						
	MergeType	OneToOne	MergeType OneToOne								OneToOne						
				-								-					
	MergeFlagVaria		MergeFlagVariab					MergeFlagVarial		MergeFlagVariab							
	RenameVariable	oldVariable: "inc	oldVariable: "inc	come"				D	oldVariable: "inc	9,	oldVariable: "inci	ome"					
			RenameVariable: "ir	ncome.y*				Kenamevariable	newVanable: *in/	RenameVariable	newVariable: "ini	come.y"					
	update	Master	update FillNew					update			FillNew						
	NewRow	FALSE	NewRow FALSE					NewRow	FALSE	NewRow	TRUE						
	KeepVariables		KeepVariables					KeepVariables		KeepVariables		1					
	DropVariables		DropVariables	+				DropVariables		DropVariables	.	1					
	MergeByNames	1	MergeByNames rowID, famID2					MergeByNames		MergeByNames	rowID, famID2	1					
Left Join		_				<u> </u>	Outer Join										
	5 Mil		t ID	I Miles		f:ID0		F 11		ćID				·			
caseID	firstName		famID rowID			famID2						lastName		famID2			
	1 Abe	15000		1 Lincoln	31000			Abe	15000			Lincoln	31000	1			
	2 Barry	NaN	11 NaN			NaN		Barry	NaN			NaN		NaN			
3	3 Chuck	17000	12 NaN	NaN Na	laN	NaN	3	Chuck	17000	12	NaN	NaN	NaN	NaN			
	Dave	25000				NaN		Dave	25000			NaN		NaN			
	Ed .	21000				NaN		Ed	21000			NaN		NaN			
	5 E0	21000	14 NaN	Nan Na	an	Nan											
										NaN		Madison	32000	10			
							NaN		NaN	NaN	3	Nelson	33000	1:			
							NaN	NaN	NaN	NaN	6	Quincy	34000	1-			
Python	DCmarant mass	no/DCmarana Jaff	_on=["caseID", "famID"], right_on=	ConstDT SeculD2T h			Duther	DCmaraat mara	o/DCmarao2 laft	_on=["caseID", "far	mIDM sight on-P	Secusion Meanings	1 harristantast)				
	Dollierge r.merg	ge(DOINEIGEZ, IEIT	_on=[caseID , IaIIID], IIgnt_on=	LIOWID , IBINIDZ'], NO	iow- ieit)			Domerge r.merg	e (υδιτιθίμας, 18π	_on=[caseib , Tar	וויים j, iigiii_on=[TOWID , IBINIDZ	j, now- outer')				
SDTL							SDTL										
MergeDatasets							MergeDatasets										
	MergeByVariabl	caseID, famID						MergeByVariabl	caseID, famID								
	FirstVariable							FirstVariable									
	LastVariable			1				LastVariable				1					
	MergeFileDescri		MergeFileDescription					MergeFileDescri	ption	MergeFileDescrip							
	FileName	DSmerge1.sav	FileName DSmerge2.sav					FileName	DSmerge1.sav	FileName	DSmerge2.sav						
	MergeType	OneToOne	MergeType OneToOne					MergeType	OneToOne	MergeType	OneToOne						
	MergeFlagVaria	DIE	MergeFlagVariable					MergeFlagVaria	J	MergeFlagVariab	ле Т	1					
	D	oldVariable: "inc	oldVariable: "ind RenameVariable: newVariable: "ir	come"				D	oldVariable: *ino	RenameVariable	oldVariable: "inci	ome"					
				ncome.y								come.y					
	update	Master	update FillNew					update			FillNew						
	NewRow	TRUE	NewRow FALSE					NewRow	TRUE	NewRow	TRUE						
	KeepVariables		KeepVariables					KeepVariables		KeepVariables							
	DropVariables		DropVariables					DropVariables		DropVariables							
			Diopvariables	-						Diopvariables							
	MergeByNames	1	MergeByNames rowID, famID2	1				MergeByNames		MergeByNames	rowID, tamID2	1					
																	

				+							
				\bot							
				\perp							
				1							
				1							
				+							

			1						
			1						
			1						

			1						
			1						
			1						
			1						
			1						

			1						
			1						
			1						

				1						
				1						
				1						
				l l						
				l						
				1						
				1						
				1						
				l l						
				1						
				1						
				l						
				I						
				1						
				1						
				I						
				l						
				I						
				l						
				l l						
				l l						
				l l						
				1						
				I						
				I						
				l						
				l						
				l						
				l						
				I						
				l						
				1						
				1						
				l						
				I						
				I						
				1						

				1							
				-							
				-							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				-							
				-							
				-							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							
				1							

			1						
			1						
			1						

			1						
			1						
			1						
			1						

			\perp							
			\perp							

				_								
				_								
				- 1								
				- 1								
				_								
				_								
				_								
				- 1								
				- 1								
				- 1								
				- 1								
				- 1								
				- 1								

Index Merge, Inn												
caseID			famID	rowID	lastName	income y	famID2					
	Abe	15000	11		Lincoln	3100		11				
	Barry	NaN	11		Madison	3200		12				
	Chuck	17000	12		Nelson	3300		13				
4	Dave	25000	13	6	Quincy	3400	00	14				
ython	left.merge(right, I	eft_index=True, ri	ght_index=True)					left.merge(right,	left_index=True,	right_index=True	, how="right")	
	3.1.0.7											
DTL								SDTL				
1ergeDatasets								MergeDatasets				
	Manage D. Afradabila	_						iviergeDatasets	Manual Di Mandala	l		
	MergeByVariable	25							MergeByVariab	ies		
	FirstVariable								FirstVariable			
	LastVariable								LastVariable			
	MergeFileDescrip	ption	MergeFileDescri	ption					MergeFileDescr	ription	MergeFileDesc	cription
1	FileName	DSmerge1.sav	FileName	DSmerge2.sav					FileName	DSmerge1.sav	FileName	DSmerge2.sav
1	MergeType	Sequential		Sequential					MergeType	Sequential	MergeType	Sequential
	MergeFlagVariab	,	MergeFlagVarial		1				MergeFlagVaria		MergeFlagVaria	
	RenameVariable		RenameVariable						RenameVariable		RenameVariab	
			update							+		
	update	Master		FillNew					update	Master	update	FillNew
	NewRow	FALSE	NewRow	FALSE					NewRow	FALSE	NewRow	TRUE
	KeepVariables		KeepVariables						KeepVariables		KeepVariables	
ļ	DropVariables		DropVariables						DropVariables		DropVariables	
	MergeByNames		MergeByNames						MergeByNames	5	MergeByName	·s
lov Morgo Lo	eft Join / Outer Jo	sin.										
			famID	rowID	lastName	income y	famID2					
	Abe	15000	11		Lincoln	3100		11				
		NaN			Madison			12				
			11									
						3200						
3	Chuck	17000	12	3	Nelson	3300	00	13				
3	Dave	17000 25000	12 13	3			00					
3		17000	13	3	Nelson	3300	00	13				
3	Dave	17000 25000	13	3	Nelson Quincy	3300 3400	00	13				
3 4 5	Dave	17000 25000 21000	13 14	NaN	Nelson Quincy	3300 3400	00	13	left_index=True,	right_index=True	e, how="outer")	
3 4 5	Dave Ed	17000 25000 21000	13 14	NaN	Nelson Quincy	3300 3400	00	13	left_index=True,	right_index=True	e, how="outer")	
3 4 5	Dave Ed	17000 25000 21000	13 14	NaN	Nelson Quincy	3300 3400	00	13	left_index=True,	right_index=True	e, how="outer")	
3 4 5 ython	Dave Ed	17000 25000 21000	13 14	NaN	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	left_index=True,	right_index=True	how="outer")	
3 4 5 sython	Dave Ed	17000 25000 21000	13 14	NaN	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	left_index=True,	right_index=True	e, how="outer")	
3 4 5 Python DTL MergeDatasets	Dave Ed left.merge(right, I	17000 25000 21000 eft_index=True, ri	13 14	NaN	Nelson Quincy	3300 3400	00	13 14 left.merge(right,			e, how="outer")	
3 4 5 5 ython	Dave Ed left.merge(right, I	17000 25000 21000 eft_index=True, ri	13 14	NaN	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	MergeByVariab		b, how="outer")	
3 4 5 bython	Dave Ed left.merge(right, l MergeByVariable FirstVariable	17000 25000 21000 eft_index=True, ri	13 14	NaN	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	MergeByVariab FirstVariable		b, how="outer")	
3 4 5 Python SDTL MergeDatasets	Dave Ed left.merge(right, I MergeByVariable FirstVariable LastVariable	17000 25000 21000 eft_index=True, ri	13 14 ght_index=True,	NaN how="left")	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	MergeByVariab FirstVariable LastVariable	les		
3 4 5 Python SDTL MergeDatasets	Dave Ed left.merge(right, l MergeByVariable FirstVariable	17000 25000 21000 eft_index=True, ri	13 14	NaN how="left")	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	MergeByVariab FirstVariable LastVariable MergeFileDescr	les	MergeFileDesc	ription
3 4 5 Python SDTL MergeDatasets	Dave Ed left.merge(right, I MergeByVariable FirstVariable LastVariable MergeFileDescrip	17000 25000 21000 eft_index=True, ri	13 14 ght_index=True, MergeFileDescri	NaN how="left")	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	MergeByVariab FirstVariable LastVariable	les	MergeFileDesc	ription DSmerge2.sav
3 4 5 Python SDTL MergeDatasets	Dave Ed left.merge(right, I MergeByVariable FirstVariable LastVariable MergeFileDescrip	17000 25000 21000 21000 eft_index=True, ri	13 14 ght_index=True, MergeFileDescri FileName	NaN how="left") ption DSmerge2.sav	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	MergeByVariab FirstVariable LastVariable MergeFileDescr	les ription DSmerge1.sav	MergeFileDesc	
3 4 5 Python SDTL MergeDatasets	Dave Ed left.merge(right, I left.merge(right, I MergeByVariable FirstVariable LastVariable MergeFileDescri FileName MergeType	17000 25000 21000 eft_index=True, ri	13 14 ght_index=True, MergeFileDescri FileName MergeType	NaN how="left") ption DSmerge2.sav Sequential	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	MergeByVariab FirstVariable LastVariable MergeFileDesci FileName MergeType	ription DSmerge1.sav Sequential	MergeFileDesc FileName MergeType	DSmerge2.sa Sequential
3 4 5 5 Spython SDTL MergeDatasets	Dave Ed left.merge(right, I left.merge(right, I left.merge(right, I left.mergeByVariable FirstVariable LastVariable MergeFileDescrip FileName MergeType MergeFlagVariable	17000 25000 21000 eft_index=True, ri	13 14 ght_index=True, MergeFileDescri FileName MergeType MergeFilagVarial	NaN Naw="left") Disnerge2.sav Sequential	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	MergeByVariab FirstVariable LastVariable MergeFileDescr FileName MergeType MergeFlagVaria	iption DSmerge1.sav Sequential	MergeFileDesc FileName MergeType MergeFlagVaria	DSmerge2.sa Sequential
3 4 5 Sython DTL tergeDatasets	Dave Ed Ieft.merge(right, i MergeByVariable FirstVariable LastVariable MergeFileDescri; FileName MergeType MergeTypa MergeTypa RenameVariable	17000 25000 21000 eft_index=True, ri est_index=True, ri DSmerge1.sav	13 14 ght_index=True, MergeFileDescri FileName MergeFiype MergeFigVarial RenameVariable	NaN Naw="left") ption DSmerge2.sav Sequential	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	MergeByVariab FirstVariable LastVariable MergeFileDescr FileName MergeType MergeFlagVaria RenameVariable	les DSmerge1.sav Sequential	MergeFileDesc FileName MergeType MergeFlagVaria RenameVariab	DSmerge2.sa Sequential ab
3 4 5 5 Yython	Dave Ed left.merge(right, I left.merge(right, I MergeByVariable LastVariable LastVariable MergeFileDescrip FileName MergeFilePascrip MergeFilePascrip MergeFilePascrip MergeFileVariable update	17000 25000 21000 eft_index=True, ri 25 25 25 25 Master Master	13 14 14 ght_index=True, MergeFileDescri FileName MergeType MergeFlagVarial RenameVariable update	NaN Now='left') ption D5merge2.sav Sequential	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	MergeByVariab FirstVariable LastVariable MergeFileDesor FileName MergeFlagVaria RenameVariable update	les DSmerge1.sav Sequential	MergeFileDesc FileName MergeType MergeFlagVaria RenameVariab update	DSmerge2.sa Sequential ab le: FillNew
3 4 5 5 Sython Storic Laboratoria Storic Laboratori	Dave Ed Ieft.merge(right, i MergeByVariable FirstVariable MergeFileDescrig FileName MergeFileDescrig FileName MergeFilagVariab RenameVariable update NewRow	17000 25000 21000 eff_index=True, ri 25 25 25 25 25 26 26 27 28 29 29 20 20 20 20 20 20 20 20 20 20 20 20 20	13 14 14 15 16 16 16 16 17 16 16 16 16 16 16 16 16 16 16 16 16 16	NaN Naw="left") ption DSmerge2.sav Sequential	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	MergeByVariab FirstVariable LastVariable MergeFileDesor FileName MergeType MergeFilagVaria RenameVariable update NewRow	les DSmerge1.sav Sequential	MergeFileDesc FileName MergeType MergeFlagVaria update NewRow	DSmerge2.s Sequential ab
3 4 5 5 Python SDTL Werge Datasets	Dave Ed left.merge(right, I left.merge(right, I MergeByVariable LastVariable LastVariable MergeFileDescrip FileName MergeFilePascrip MergeFilePascrip MergeFilePascrip MergeFileVariable update	17000 25000 21000 eft_index=True, ri 25 25 25 25 Master Master	13 14 14 ght_index=True, MergeFileDescri FileName MergeType MergeFlagVarial RenameVariable update	NaN Now='left') ption D5merge2.sav Sequential	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	MergeByVariab FirstVariable LastVariable MergeFileDesor FileName MergeFlagVaria RenameVariable update	les DSmerge1.sav Sequential	MergeFileDesc FileName MergeType MergeFlagVaria RenameVariab update	DSmerge2 Sequentia ab le: FillNew
3 4 5 5 Python	Dave Ed Ieft.merge(right, i MergeByVariable FirstVariable MergeFileDescrig FileName MergeFileDescrig FileName MergeFilagVariab RenameVariable update NewRow	17000 25000 21000 21000 eff_index=True, ri ss ss ss Master TRUE	13 14 14 15 16 16 16 16 17 16 16 16 16 16 16 16 16 16 16 16 16 16	NaN Now='left') ption D5merge2.sav Sequential	Nelson Quincy	3300 3400	00	13 14 left.merge(right,	MergeByVariab FirstVariable LastVariable MergeFileDesor FileName MergeType MergeFilagVaria RenameVariable update NewRow	les DSmerge1.sav Sequential	MergeFileDesc FileName MergeType MergeFlagVaria update NewRow	DSmerge2. Sequential ab le FillNew