Merging Datasets in SPSS

Sometimes you'll need to merge datasets together in SPSS. *In order to merge two data files, both files must be saved as .sav (SPSS data) format*.

1. You will need to pick a "key variable" on which the datasets will be matched. *You must sort both datafiles ascending by the key variable and then save them*.

🔓 Untitled4 [DataSet3] - IBM SPSS Statistics Data Editor											
<u>F</u> ile	<u>E</u> dit	View	<u>D</u> ata	<u>T</u> ransform	<u>A</u> nalyze	Direct <u>M</u> arketing	<u>G</u> raphs	<u>U</u> tilities	Add- <u>o</u> ns		
			☑ Define <u>V</u> ariable Properties ☑ Set Measurement <u>L</u> evel for Unknown								
		V		opy Data Pro			var	var			
	1		N	lew Custom A	ttri <u>b</u> ute		vai	vai			
	2	1	🖶 D	efine Dates							
	3	1	🔡 D	efine <u>M</u> ultiple	Respons	e Sets					
	4	1	V	alidation		4					
	5	1	🔡 Id	lentify D <u>u</u> plica	ite Cases.						
	6		🗔 <u>I</u> d	lentify Unusua	al Cases						
	7		🔊 s	ort Cases							
	8		S S	ort Variables.							
	9		TI TI	ranspose							
	10		1.4.941	erge Files		•	Add	Cases	<u> </u>		
	11	-		estructure				-	_		
	12	-		ggregate			+ Add	l <u>V</u> ariables	•		
	13	1		rthogonal De	eion	•					
	14			-	argii	, , , , , , , , , , , , , , , , , , ,	_				
	16	1		opy <u>D</u> ataset			_				
	17	1		plit <u>F</u> ile							
	18	1		elect Cases							
1	19	1	4 9 <u>N</u>	eight Cases.							
	20	1									

2. Select "Data > Merge Files > Add Variables" from SPSS's main menu.

3. Tell SPSS that you want to merge variables from an external file:

🙀 Add Variables to Study 1 Demographics.sav[DataSet8]	×
Select a dataset from the list of open datasets or from a file to merge with the active dataset \bigcirc An $\underline{o}pen$ dataset	
Untitled4[DataSet3]	
An external SPSS Statistics data file	
sional\Classes\PSYC 6384 - IO Methods\Data\Week 1\Study 1\Study 1 Manager Ratings.sav	Browse
Non-SPSS Statistics data files must be opened in SPSS Statistics before they can be used as part	of a merge.
Continue Cancel Help	

4. Set the merge options appropriately:

	av X					
	Excluded Variables: New Active Dataset					
Check			*	gender_Male(*) gender_Female(*) gender_Trans(*) gender_Other(*) age(*) race_Asian(*) race_Black(*) race_caucasian_Whit		
this box		Rename		race_Hispanic_Latino(*)	This is the	
	Match cases on key variables in sorted fi			Key <u>V</u> ariables:		
	O <u>B</u> oth files provide cases		*	employee_id	variable the	
	Non-active dataset is <u>keyed table</u> Active dataset is keyed table		-		files will	
	Active dataset is keyed table					
					match on	
This opt	ion means that the	ofessional\Classes\PSYC 6384 - IO Methods\Data\Week 1\Study 1\Study 1 Manager Ratings.sav		I		
file you'r	e merging can have	OK Paste Reset Cancel Help				
only	<i>one row</i> per key					
(mat	tching) variable					

When matching two datasets, one file is called the *keyed table*. The primary dataset can have multiple rows per participant (e.g., five rows for participant #3). However, the *keyed table* can have only one row per participant (e.g., one row for participant #3). All of the variables in the *keyed table* will be appended to the matching rows in the primary dataset (e.g., all of the variables in the row for participant #3 in the *keyed table* will be appended to the end of all five rows for participant #3 in the primary dataset).

- 5. Click "Okay" and enjoy the merged dataset!
 - Note that the merge <u>will fail</u> if both datasets are not sorted ascending on the key (see step #1) or if the keyed table has more than one row for any participant.