Handling Large-Scale Unit Level Data Using STATA Professor Pratap C. Mohanty Department of Humanities and Social Sciences Indian Institute of Technology, Roorkee Lecture 39 Construction of Panel Data

Welcome once again to the NPTEL module on handling large scale data, using STATA. We are now explaining the panel data and their construction, how we can able to construct the panel data correctly.

(Refer Slide Time: 00:39)



And that will help us to go for panel model and the analysis very systematically and correctly. Coming to the understanding of panel data. Once again, we just wanted to emphasize that panel data has observation on the same units in several time periods. Many times the panel data is not provided in ready to use manner, and we have to construct them in a panel format.

The separate data file that is provided for each year is important. We have to merge them for our analysis. In India particularly, the only longitudinal data in the panel format is through India human development survey datasets. So, IHDS is very important that is unique in this particular direction. And we are just clarifying IHDS data once again though we have done it earlier, but in a panel setup we did not mention earlier.

(Refer Slide Time: 02:01)



I just wanted to mention once again that IHDS is a national representative data with multi panel survey of households across India. IHDS is jointly organized by University of Maryland. And the NCAER that is the National Council of Applied Economic Research based in New Delhi. IHDS public data files are available from the Data sharing and Demographic Research of inter University Consortium for Political and Social Research. That is ICPSR, I will show you all those links.

(Refer Slide Time: 02:37)



Here is the link ICPSR link once you click on this link, you will be directed to the exact website where the data is available. So, IHDS is the first large scale national panel survey of over 40 thousand Indian rural as well as urban households undertaken by the researchers from NCAER and University of Maryland. There are in fact two rounds of the survey and those have been conducted since 2004-05 till 2011-12. So, 2004-05 is famously known as IHDS one and second that is 2011-12 is called IHDS 2.

(Refer Slide Time: 03:18)



IHDS-1, surveyed over 41554 households. And these numbers are very important. While understanding panel, we need to remember these numbers because then only you can able to understand how many are dropped and how many are exactly matching. This is very very essential while constructing panel data. So, out of 41554 households, 26734 are from rural and 14820 are from urban households.

Total number of household, surveyed in IHDS-2 is of 42152 and those consist of again 27579 rural and 14573 urban households. IHDS-2 re-interviewed that is important to mention re-interviewed. We are going to clarify 83 percent of the original households. So, in the second wave, that is IHDS-2 considered 83 percent of the original that is IHDS-1 data as well as split households residing within the household and additional sample of 2134 households' additional samples.

We are going to clarify everything. This is the number we mentioned and IHDS-2. It is of 42152 households surveying both round that those common is of 40018 and the household lost re-contacted in second wave and they are of 6911 and the households that are not included in the second round of the surveys of 2134. And why these are the cases we are going to clarify right now.

(Refer Slide Time: 05:32)



The questionnaires are in fact similar across the two waves to enable comparison over time. And it has been very clearly mentioned that two rounds can be compared. Why comparison important? Those who are interested in doing some experimental framework for analysis, in comparison some effectiveness of a particular policy over that two-time period, if it is somewhere suitable is going to be very useful. (Refer Slide Time: 06:03)



Understanding the linking of the both the rounds of the data in order to link two datasets, we require a linking file that is important I am underlying it so, just to emphasize that whenever two rounds of data to be merged to make a panel which one are common by the exact unit, the exact respondent or the household is the unit of the survey. Those linking files must have been available.

So, linking information is available at both households and individual level in the IHDS website. So, linking files gives round one identification codes for all the round two households that were interviewed at both the levels, that is for individual level as well as household level. You will need to register at IHDS Website to download this linking data. (Refer Slide Time: 07:06)

https://iho	ds.umd.edu/data/data-download
	INDIA HUMAN DEVELOPMENT SURVEY ABOUT THOS DATA PUBLICATIONS & RESEAR
	From External Site
	These data are freely analized on http://www.kprcumich.edu Oowelood INOS-I (2006-5) Oowelood INOS-II (2011-12)
	Linking Data Download From IHDS
	Indian Paradox: Rising Education, Declining Women's Employment
	Please sign up/login in order to download the following files. (See menu at top for link.)
	INDS Ever Married Women Linking Variables - Text File
	HIDS Ever Married Women Linking File - Stata Format HIDS Individual Linking Variables Text File
	HDS Individual Linking Variables-STATA format

So, just registration. If you do it, it is freely available. You need to give minimum information for registration. You will be given this page to download the data. It gives information such as understanding ever married women. They are in also text file, in STATA file also, individual linking variables are available. Then household linking variables are also available.

Linking data is essential sometimes, as they also mentioned, that understanding rising level of education, declining women empowerment, some paradoxical changes in two rounds are also important.

(Refer Slide Time: 07:53)



Coming to the linking files that are available in text format as well as in STATA format of 2011-12 format. If you have earlier version, you can convert it to the other version of the data also. A note here that ever married women file will be merged with the household file. it is important to mention, ever married women file will be merged with Household File since IHDS-1 does not give information specifically on ever married.

So, in that case, you have to merge with the household file and household file will be merged with certain limited information related to ever married women. Basically it is 15 to 49 age of women. Further clarifications on understanding the relationship between these two rounds for merging to make a panel data.

(Refer Slide Time: 09:05)



We first try to understand the structure of the data in both the rounds. In IHDS-1 sample, if we consider it in total, it is of 41554, and there were attrition number of observations. Those have been dropped is of 6911 may not be those observations are available. And there might have been some migration or that might have been some lapses. And there are various reasons of not covering the same individual household from the same location there might even change sifted to another place. So there is an attrition that will be subtracted from the original data.

How many replacements are there in that place in the new round that is of 2134 will clarify in our next slides in detail. There are some split households also like in the household gets splited to a number of observations like in different lead to different households, so those all have been included. That is of number 5397. Some re-interviewed households are there. This is the re-interviewed means whatever they are, out of 41000 straight away 31621 have been directly re-interviewed in the next round.

(Refer Slide Time: 10:40)



We will clarify this replacement and split household once again. Also, the attrition households are the one who are in the process get dropped out from the questioning round or the question round of the next level. Attrition it means that IHDS-1 household lost to re-contact in IHDS-2. Some of the original IHDS households had to be replaced in some urban areas where interviewers were unable to locate the former households,

So, those have been shifted or like replaced in some urban areas, very difficult to locate. They are also part of the attrition level. But a clarity is also given in the dataset related to the replacement household like a replacement household was randomly selected in the same neighbourhood to refresh the sample. This has led to 2134 new households being included in the IHDS-2 sample. So, that has been contacted to the neighbourhood and more or less the same nature of sample has been replaced. That is of 2134 new households. These are called new replacement household. So, replacement households are identified by a 9 in the split file 9 is the code entered for the replacement household.

(Refer Slide Time: 12:20)



It means there are some households which are not interviewed in IHDS-1. Basically, since they are replaced in a new survey. That is very clear that number that is of 2134 have not been included or were not including IHDS-1. So, households from IHDS-1 may have split into two or more households in IHDS-2. Like one household between IHDS-1 and IHDS-2 the number of years, from 2005 till 2011-12 around 5-6 years in between, there are high possibility of some split households.

One household gets split to two or more households. And they are all included in IHDS-2. If the split household resided in the same village or urban block. All splits were included for survey. Because of the split households in IHDS-1 match with more than one households in IHDS-2. Because of the split. So, the total number of household surveyed in IHDS-2 is what exactly that is.

(Refer Slide Time: 13:42)



N equal to 2134, plus 2134. We have said that the new replacement households, and the split household that would of 5397 plus re-interviewed households. So, they are of 34621 in total it is of 42152. So, the total observation we get in IHDS-2 but they may not be entirely of longitudinal type, entirely of final types. The exact number we are going to discuss right now.

Households surveyed in both the rounds, this is more important to be noted for our analysis that is split household from around 5397 that we have already mentioned, plus the re-interviewed that is of this. So, 34621 if you add these two. So, both the rounds and their observations are now similar and that is of 40018 number similar in both rounds.

(Refer Slide Time: 14:54)



And this number is important for panel analysis. Coming to the linking of two datasets once again to mention that IHDS-1 household in total were of 41554 and in two it is 42152 household link file contains household linking, file contains the entire 42152 households so linking file will provide IHDS-1 identification codes for all IHDS-2 households and missing codes for replacement households that we have already mentioned, as 9 is the code with the number 2134, And those are very new.

Coming to the linking variables, mark it very carefully linking variables. They are specifically state ID, district ID, psu ID, HHID, household basically the ID label. Then this is of primary sample unit they have defined then with an id this is the split id and this is person ID. So, to be noted here that the person ID is required, especially this particular ID is required when we are linking for the individual files, not for the household files.

(Refer Slide Time: 16:18)



Coming to the steps of linking IHDS and IHDS-2, we need to launch the data now onwards, for another 15 to 20, around 13-14 slides will be experimenting with the original data for linking its one and two to make a panel data. With 40000, the number of which already 40018 is the number and wanted to find out by linking. So, let us use the dataset that is IHDS-1 household data. At this moment we are linking and we will open.

But I wanted to mention here that in order to get the link very clearly it might be the case that both the datasets that is IHDS-1 household and IHDS-2 household datasets have the same name, If the same name is there but we are linking from Masters to the new both the variable name might be merged with the same name and the information with the new variable that will be identified in the mod database will carry the master's variable name.

So, what does this mean basically when it has two variables separately and to both the variables, give different information. But, we are carrying the same name that will be misleading. Alright. Since both the variables are different in different rounds and they give different information, so in order to differentiate the variables, we are renaming the variable. first of all, we rename the one dataset that is IHDS-1.

So, you need to rename the entire set of variables. Those are given in IHDS-1 household data. I suggest that you please likewise do for renaming of the variable name of a particular variable, then its new name, it gets renamed. But since we want to differentiate by all the variables with all

the variables in another round, for simplicity we take x is another indicator addition to the name of the variable.

So, the exact suggestion is that there is space mark here between the asterisks to asterisks mark. There is a space here. So, we are highlighting here there is space between first asterisks symbol and X, in between this space mark, if you will just do that, all the variables will be added with x. We are going to show it, with the IHDS-1 household data. So, it is here.

(Refer Slide Time: 19:16)

ew T 4 x hitter commands here 0							
Command or	Special Edition	4905 Lake College 5	way Drive Itation, Texas 77845 USA			Variables	¥ 0
		800-STATA 979-696-4	-PC http://www.stati 600 stata@stata.com	com		Name	Level
use "G\panel data analysis\		979-696-4	601 (fax)			xCASEID	HH ID: 11 byte stn
rename ' x'	it-student Wata lab perpet	and linear				STATED	State code
rename xSTATEID STATEID	Notal Autor: 4111	DEDRIVES #				OISTID	District code
rename x0IS1ID 0IS1ID rename xPSUID PSUID	Licensed to: IIT	Roorkee				PSUID HHID2005	PSU: village/neigh. Household ID
rename xHHID HHID2005						HHSPUTID20	L. Split household ID
rename xHHSPUTID HHSPL	lotes: 1. Unicode is suppor	ted: see help unicod	e advice.			иОнн	HH ID 9-digit uniq
	2. Maximum number of	variables is set to	5000; see help set_maxvar.			A XIDPSU	PSUID 6-digit uni.
	use "C:\panel data analys	is/IHDS1 HH.dta"				XOISTNAME	District codes with
((India Ruman Development Su	rvey (IRDS), 2005, H	iousehold)			xDIS101	H1sp: District ID C.
	rename * x*					YURBAN	Census 2001 villa.
						E XMETRO6	Largest 6 metro ar.
	rename ESTATEID STATEID					Properties	u u
						4 Variables	
	rename #DISTID DISTID					Name	HHID2005
						Label	Household ID
	rename xPSUID PSUID					Type	
						Value label	HHD
-	rename wHHID HHID2005					Notes	
						4 Data	NUMBER AND ADDRESS
	rename xERSPLITID RESPLIT	102005				Label	India Human De
						 Notes 	
						Variables	945
C.	ommand					Observations Size	41,554 100.97M
						Memory	160M
SE 151 Edit Data Graphics Statistics	Uter Exc	1 2 3	^{\$} 4 [%] 5 [^] 6 ^{&} 7	890	e Bksp	- 10(= 0
aSE ISI Edit Data Graphics Statistics V T 3 X Iter commands here 0	Uter Esc - , Tab q Caps	1 ¹⁰ 2 * 3 w e r a s d	⁸ 4 [%] 5 [°] 6 [®] 7 Ityui fghj	8 9 0 9 1 1	i liksp j liksp j liksp j liksp j liksp j liksp j liksp j liksp	 Yaiables Fiter canab 	11/12/2020
Est Data Giaphics Statistics Est Data Giaphics Statistics F T 3 × ter connands here Connand JK	Uter Ex - , Tab q Caps Shift	1 ¹⁰ 2 ¹⁰ 3 w e r a s d z x c	∫ ⁵ 4 [№] 5 [^] 6 [®] 7 t y u i f g h j v b n r	■ 8 9 0 0 p 1 7 7 k 1 : 1 . 1 . 1 . 1 . 1 . 1 . 1 . 1 . 1 .	Eissp 1 V Del Shift	 Variables Fiter variables Name 	to here Label
SEIST Total Graphics Standings Total Standings Total X er commands here animand J.C. There are so dems to show.	Uter Esc Tab q Caps Shift Static Ctrl 4	a s d Z x c	§ 4 [%] 5 [^] 6 [®] 7 t y u i f g h j v b n r	8 9 0	Bitop 1 V Det Shift T Fn	 Variables Efter variables Name 	to them to show.
St 131 Got Data Graphics Stanting Image:	Uner Esc Tab q 	a s d Z x c	§ 4 [™] 5 [°] 6 [™] 7 t y u ji f g h j v b n r ∧t	8 (9)0 0 p (1) 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Bitop Bitop J Del Shift T Fn	Variables Variables Variables Variables Variables Variables vorth panel data andy. P	In them to show.
ESSI at Data Gaptica Statistica bit Data Gaptica Statistica T 3 × 4 - 8 - 9 3 T 3 × 7 continuadi herr or 0 familio to blow.	Un Exc	a s d Z X C AR	[⁵ 4] [™] 5 [°] 6 [®] 7 [t]y]u]i [f]g [h]i [v]b]n]r µ]art		Bitop 1 Del Shift Fn	Yaiables Variables Variables Variables Variables Name Name variables variabl	Liner Label No items to those.
151 t bits capho: Standic tit tit <td>Urr Ssc - , Tab q Caps - , Salt State Sp</td> <td>1 9 2 3 w e r a 5 d z x c Alt</td> <td>\$ 4 % 5 6 % 7 t y u p f g h j v b n n At At At</td> <td>8 9 0 - 0 P 1 0 k 1 - - n c, > - - n c, > - - n c, > - -</td> <td>set</td> <td>Variables Variables Variables Kane extr porel data andy P E C 0</td> <td>ies here Label ho items to show.</td>	Urr Ssc - , Tab q Caps - , Salt State Sp	1 9 2 3 w e r a 5 d z x c Alt	\$ 4 % 5 6 % 7 t y u p f g h j v b n n At At At	8 9 0 - 0 P 1 0 k 1 - - n c, > - - n c, > - - n c, > - -	set	Variables Variables Variables Kane extr porel data andy P E C 0	ies here Label ho items to show.
1013 Graphics Statestics 1014 Graphics Statestics 1014 Graphics Statestics 1014 Graphics Statestics 1014 To a x	Un Fie Con	a s d z x c Alt Cognoze New folder	3 4 % 5 6 % 7 1 1 1 1 f 9 h j v b n n At At	8 9 0 . 0 P 1 P k 1 2 7 % m <.p>.p 2 7 % m c 2 7 % m c 2 7 % m c 1 2 7 % m c 1 2 7 % m c 1 2 7 % m c 1 1 1 1 1	Shift Shift	Variables	ies here Label to items to those.
SSI t the Graphic the Gr	Corr 20 State 20 Corr 20 Corr 20 C	1 2 3 W e r a s d 2 X c Art Cogarae · New Islder Doctrop is Doctrop	Image: Second	A 9 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	State 2017 KS	 Yarables Yarables	In N LUZZON
351 It Data Gupton Statesco It Data Gupton State	Conference of the second secon	1 2 3 W e r a s d z X c Alt Cognite - New Noter Fonoties Dominach Spicent Places	Image: second	8 9 0 0 p 1 0 p 1 1 cori 1 0 cori 1	Ser 2.24 KB 2.24 KB 2.	Variables	L N LUXAN
Bai Casho Somma Casho Somma T P	Correction of the second secon	1 2 3 3 W e r a 5 d 2 X C AR Cognite - New folder Destrop Destrop Destrop Destrop Destrop Destrop	5 4 % 5 6 6 7 1 1/	8 9 0 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 1 0 0 0 1	Sear 2.244 48 2.207 4	Variables Variables Effer sand Name	C
(33) stringer (34) - ≤ - ≤ - ≤ - ≤ - ≤ - ≤ - ≤ - ≤ - ≤ -	License in pro-	1	Image: A manual strain strai	8 9 0 1 0 P 1 1 k 1 2 1 iii crit 1 1 iii crit 1 1 iiii crit 1 1 iiii crit 1 1 iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii	501 501 501 501 501 501 501 501 501 501	Variables	C C C C
IB3 IM Data Graphic Statistics	Arrow Carlos Car	1 2 3 W e r a 5 d 2 x c AR Control Dentrol	3 4 % 5 6 6 7 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	B 9 0 1 0 0 0 0 0 k 0 2 0 0 m c 2 2 1 m c 2 2 1 m c 1 2 1 m c 1 2 1 m c 1 2 1 m c 1 1 1 m c 1 1 1 1 m c 1 1 1 1 1 m c 1	5cm 2.34 K3 2.07 K8 2.34 K3 2.07 K8 105/9743 105/9743 105/9743	Variables Variab	C C
BD Capton Stanting Capton Stanting Capton Stanting T a k Command by T a k Command by T a k Capton Stations Capton Station Capton Stations Capton Stations Capton Statio	Correction of the second secon	1 0 2 8 0 a 5 d 2 X C AZ X C AX D	1 4 5 6 0 7 1 1 1 1 1 1 1 1 1 1 1 1 1 1	8 9 0 1 0 P 1 1 1 1 Crit - - 1 1 1 Crit - - - 1 1 1 Crit - - - - - 1	5ce 2.374 4.3 2.374 8.3 1.2376 8. 1.2376	Yauldes Yauldes Yauldes Yauldes Tore reading Tore reading Tore and any. P S	C
ESS et Das Gapho Sontho B C S + δ + df dS et S + δ + df dS et S + δ + df dS et S + df dS et	Capa Capa	1 0 2 3 3 2 X C 7 3 2 X C 7 7 3 2 X C 7 7 4 2 X C 7	Image: Section 1 Image: Section 1<	8 9 0 0 P 0 1 0 P 0 1 0 P 0 1 0 P 0 1 0 P 0 1 0 P 0 1 0 P 0 1 0 P 0 1 0 P 0 1 0 P 0 1 0 P 0 1 0 P 0 1 0 P 0 1 0 P 0 1 P 0 0 1 P 0 0 1 P 0 0 1 P 0 0 1 P 0 0 1 P 0 0	See 2,544,83 2,524,84 2,524,84 2,527,547,547,547,547,547,547,547,547,547,54	Vesibles Vesibles Three Ref apprend and any. P	The NUMBER OF STREET
ESS et Das Gapho Sontos Sontos V Sontos T is connuad reg 0 neural R 0 R 1 R 1 R 1 R 1 R 1 R 1 R 1 R 1	Or Cap	A Control Contro Control Control Control Control Control	1 4 5 6 0 7 1 1 1 1 1 1 1 1 1 1 1 1 1 1	B 9 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0	Bits Fr 9-10 - 9-11 -	Vaulles Vaulles Nor Nor Nor Nor Nor Nor Nor Nor Nor Nor	C C C C
ESS et Das Gupho Sontion to the Company of the second to the second base error are to ben't to box. Sontial Second Secon	Correction of the second secon	A Coprate	1 4 5 6 0 7 1 1 1 1 1 1 1 1 1 1 1 1 1 1	B 9 0 I P I I Crit I II Crit I III Crit I III Crit I IIII Crit I IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	5ar 2.344 43 2.327 48 2.327 48 2.327 48 10.529	Values	The NUMBER
ABI et Data Graphica Standing B B - df - df - df - df B B - df - df - df Contracting the bare 0 extracting the bare 0 here are no forms to above. 0 a 0 a 0 a 0 a 0 b 0 b 0 a 0 b 0 b 0	Capa Capa	A Constant Consta	Image: A manual stress and the stress and t	B 9 0 0 P 0 0 P 0 10 Col 0	Bitop Shit Per Shit Fm 2,548,48 Fm 2,549,48 Fm	Validis	Lord Lord
XESS Left Das Gaphen Stender Life Tar And State Life Tar And State Variation Named	Or Cap	A Compare	1 4 5 6 0 7 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	B 9 0 I Crit I I I Crit I I I I Crit I I I I I Crit I I I I I I Crit I	Bits Fr • Fr	Variables Variables	T R INDEXES
All bit Capiton Stanting If Dials Gaption Stanting If If I are dependent of the stanting	Correction of the second secon	Control Control Contro Contro Control Control Control Control Co	1 4 5 6 0 7 1 1 1 1 1 1 1 1 1 1 1 1 1	B 9 0 I P I I Crit I II Crit I III Crit I III Crit I IIII Crit I IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	5ar 2.34 4.3 2.32 4.4 2.32 2.4 12.32 4.4 12.32 4.4 12.32 4.4 12.32 4.4 13.32 4.4 13.34 4.4 14.34 4.44 4.44 4.44 4.44 4.44 4.44 4.44	Values	T R INTERNE
AB131 Capton Standard If Data Gapton Standard If If I and I a	A Capa	0 2 3 3 0 7 7 a 5 d 2 X C 7	Image: Section 1 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2	B 9 0 0 P 0 0 Crit 0 10 Crit 0	Bilop Shit Per Shit Fm 2,548,48 Per 2,549,49 Per	Validis	T R INTERNE
OTINI (eff) Out 6 Capton Standow (f) Out 6 Capton Standow (f) Out 6 Capton Standow (f) Out 7 Capton Standow (f) Out 7 Capton Out 7 Capt	Or Cap Cap Later of the second	A Constant of the second	1 4 5 6 0 7 1 1 1 1 1 1 1 1 1 1 1 1 1 1	B 9 0 I Crit - - I Crit - - - I Crit - - - - I Crit - - - - - I Crit -	5cm 2.307.48.4 2.307.48.4 2.307.48.4 10.104.48.48.4 10.104.48.48.4 10.104.48.48.48.48.48.48.48.48.48.48.48.48.48	Variables	To Re 11/12/2008
AB131 Left Data Graphic Standard I III - A - A - A - A - A - A - A - A -	Capa Capa	Curran C	Image: Section 2010	B 9 0 P 0 0 R 0 0 0 R Crit - 0 B Crit -	500 500 500 500 500 500 500 500 500 500	Validis Validis Ture Ture Constant for many E - 0	T R IVIDAD
RED3 Gapta Series If Data Gapta Series If Series If Series If Series If Series If Series If Series If Series If Series If Series If Series	A Capacity of the second secon	A Constant Constan	1 4 5 6 0 7 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1	8 9 0 I Crit I I I Crit I I I I Crit I I I I I Crit I I I I I I Crit I	5ar 23748 23	Variables Variables	E R 12/22/00
RED3 Coppen Statute If Das Gapten Statute If I I I I I I I I I I I I I I I I I I	Constant from the second secon	Coperative Control of the contr	1 4 5 6 0 7 1 1 1 1 1 1 1 1 1 1 1 1 1	B 9 0 k P 0	5ar 2.34 8.8 2.327 8.8 100,599.8 10,199	Vasidis Vasidis Vasidis Vasidis Torrector tore	In the LULEYONG CONTRACTOR OF
ACT 131 feet (the Copies Graphen Strettens (the Copies Strettens (the Copies Strettens (the Copies Stretens	Verify and the set of	Constant	Image: Section 1 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Name Example, addressed data Image: Section 2 Image: Section 2 Name Example, addressed data Image: Section 2 Image: Section 2 Name Example, addressed data Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: Section 2 Image: S	B 9 0 Crit	5 srit	Validis Validis Tare Tare Tare Tare Tare	Constant of the second se
COUNT Ent Couples Section Ent Couples Section The Couples Section Conservation Ent Couples Section Ent Couples Ent Coup	Orread	A Carper	1 4 5 6 0 7 1 1 1 1 1 1 1 1 1 1 1 1 1 1	B 9 0 I Crit I I I Crit I I I I Crit I I I I I Crit I I I I I I Crit I	5 cm 2 201 48 2 201 48	Variables Variables	C C



So, we are going to open it. It is here. IHDS-1 household data this is the one, isn't it? And so we will rename. Then asterisks say space is there then x and with asterisk enter. You just mark all the variables having added with x, x was not there earlier. All the variables are added with x before the variable name you just mark very carefully. X is there in every variable alright. Coming to the same interpretation.

This is what we have clarified, this command adds a prefix of x to all variables. Renaming is important because when we merge two files, it will overwrite the variables. So, if you want to keep variables from wave one, that is IHDS-1 and we have to separate. You want to just keep all the variables separate. You may need to rename all the variables and we have already renamed.

(Refer Slide Time: 20:34)



So, rename the id variables as original. One id variable to be original then only based on the id variable. Such as the similar id variable which are based on that. The same id will be carried to another one and that same information against the individual id will be added with different variables, different information. So, ID variables should be the same. What we will do? We will change the ID variable. We have now changed all ID variables to be added prefix with x, x is there. We will change to only the same name that is state ID.

Similarly, district ID, PSUID but another information to be very important that the household ID. Since we have changed the variable name of Particularly 2005. Our household ID is added with x but we wanted to mark that it is of 2005 households because in split ID we are going to the split folder, split file the variable name with HHID 2005 is given also a split household 2005 is given.

So, we have to make the renamed the household file to be HHID household ID with 2005. So, if you do that, simply renaming this we will get those name, which is going to be very useful. So, HHID and household split ID is not same in both the waves. Which I just said now. In linking file given with name HHID 2005. When HHID split ID2005. So, renamed it as per requirement. Since it is essential we will rename it. So, you can follow this step on your own and it is very easy to do alright. So, like we can do it so rename we will do it here. Alright. We will get to that clearly, so it is here, we will simply rename all those files. Alright.

(Refer Slide Time: 23:00)

eview T # x		4905 Lakeway Drive	Variables	T 0 >
hiter commands here	Special Edition	College Station, Texas 77845 USA	+ Filter variable	
Command		800-STATA-PC http://www.stata.com	Name	l raiar
use 'Glassel data analysid	-	979-696-4601 (fax)	v ASEID	HH ID 11 bute sto
rename i vi			STATED	State rode
resame (STATEID STATEID	25 student Stata Lab perpetu	A LINNER	DISTID	District code
CONTRACTOR ON THE OWNER	Berral Butters 47750	Conce.	8500	Kit ullaselseich
rename vBCUID PSUID	Licensed to: 11T p	orxee	HH02005	Hourshold ID
include a station station			HILDOOD HILDOOD	Fridaktion to
comme annual AMU/2003	Notes:		HESPOTION.	NU ID 0 diaturia
Tenanie Annoru tu nhoru-	 Unicode is support. 	d; see help unicode_advice.	1044	RH 10 9-0 git uniq.
	2. Maximum number of	variables is set to 5000; see help set maxvar.	h MUYSO	PSUID 6-algitum_
			PSTATED2	State codes, colu
	. use "G:\panel data analysi	NINDSI_NH.dta"	LUSINAME	Danct codes with
	(india Waman Development Sur-	ey (ikus), zuus, kousesoid)	¥JI5101	H1sp: District ID C
	. rename * x*		N.N.SAN	Census 2001 vita.
			XV02180,00	Largest 6 metro ar
	. rename xSTATEID STATEID		Properties	ų.
	VARIANA POTORID DISTIN		 Variables 	
			Name	HHID2005
			(abe)	Household ID
	. rename xPSUID PSUID		Type	int
			Format	%17.0g
			Value label	
	. rename skeid skidzoos		Notes	
			4 Data	
	. rename #RESPLITID RESPLITIO	2005	Filename	UHDS1 HHLdta
			Label	India Human Deve
			 Notes 	
			Vanables	945
	Command		a Observations	41,554
			Size	100.97M
				160M
			Sorted by	
Stata15				CAP NUM C

Simply, if you have renamed those ID, specifically the common ID which we wanted to discuss that state ID, district ID, PSUID has been removed with x and others are also removed with x and we added 2005 against to this HHID 2005 and HHID split ID 2005 rest variables we have kept it with x as the prefix. That is one of the guidance, have done it.

We have very clearly marked the variables. Those are common, we will sort those as for merging because there are two rounds. In any case, we have to merge with their common ID. So, this common ID which we have already defined and it has been already instructed in IHDS rounds also which are the common ID for both the periods and for individual periods.

So, since these are the common ID, we will sort first. As per the merging guidance we have already given in earlier lectures. So, we will sort it. We will simply sort those five IDs state ID these, this, this and this, and we will save and replace. Alright. So, with a name, I think it has already been saved.

(Refer Slide Time: 24:46)



So, it is with IHDS-1 alright.

(Refer Slide Time: 24:47)

Review T # X	800-STATA-PC http://www.stata.com	 Variables 	÷.
 Filter commands here 	979-696-4601 (fax)	* Filter variabl	es here
# Command Jrc		Name	Label
use 'G\panel data analysis\	Hata Lab perpetual License	xCASEID	HH ID: 11 byte stri
2 rename * x*	tal namer withstate	STATEID	State code
rename xSTATEID STATEID	icensed to: 111 Koorkee	OIS 11D	District code
rename x0ISTID DISTID		PSUID	PSU: village/heigh
rename xPSUID PSUID Notes:		HHID2005	Household ID
rename xHHID HHID2005 1.	Unicode is supported; see help unicode_advice.	▲ CHSPLITID 20	05 plit household ID
rename xHHSPLITID HHSPL 2.	Maximum number of variables is set to 5000; see help set_naxvar.	XIDHH	HH ID 9-digit uniq.
sort STATEID DISTID PSUID.		xIDPSU	PSUID 6-digit uni
. use "G: (India Re	panel data analysis\impi_mm.dta" an Development Durbay (INDS) 2005 Moutebold)	xSTATEID2	State codes, colla
Consta and	an pereropent server (into), rees, economy	XOISTNAME	District codes with.
. rename	x.	xOIST01	H1sp: District ID C.
		xURBAN	Census 2001 villa.
. rename	STATEID STATEID	xMETRO6	Largest 6 metro ar.
		E Properties	a a
. zenane	DISTID DISTID	· · · ·	
		 Variables 	
	PSUID PSUID	Name	HHSPLITID2005
		Label	Split household
		Type	int .
. rename	SHID SHID2005	Format	%12.0g
		Value label	HHSPLITID
·		Notes	
. Periane	MRSPEITED MRSPEITEDVOS	4 Data	
. sort ST	TEID DISTID PSUID MHID2005 MHSPLITID2005	Filename	IHO51_HHU8ta
		Label	India Human De
		 Notes 	
		Variables	945
Command		a Observations	41,554
save hh il		Sae	100.97M
		Memory	160M
		Sorted by	

So, you have already saved. So, we come to the guidance once again.

(Refer Slide Time: 24:54)



Coming to another step, the first step related to sorting, the important common IDs we have made. Now will process it for the IHDS-2 household. Similarly, will open IHDS-2 household data because now we are merging the household data. So, in this case, again, let me open that first. So, let me clear it here. So, where is clear.

(Refer Slide Time: 25:23)

ile Edit Data Graphics Stati	tics User Window Help		
	😫 🖪 O · O		
view T a -		 Variables 	τ 8
Edu commente have	Notes:	A Comment	las here
Tiner commands nere	 Unicode is supported: see help unicode_advice. 	riter varias	
Command	Maximum number of variables is set to 5000; see help set maxvar.	Name	Label
use 'G\panel data analysis\	. use "G:\panel data analysis\IND51 NH.dta"	SURVEY	0H05-1 (2005) of L.
rename " x"	(India Human Development Survey (IHDS), 2005, Household)	STATED	HQ0 State code
rename ISTATED STATED		UDID	HQU District code
rename kuls nu uls nu	. rename * x*	PSUID	HQUPSU: Wrage/
rename xPS0ID PS0ID	rename x5TATEID STATEID	HHLD	HQ0 Household L.
rename kning/ nng/2003		Instanto	Provi spiri nouser.
rename infrarunto miseu.		10430	Psu la unique e un
son shireb bishb rsolb.	. rename xDISTID DISTID	WT.	Weight is SIDE T
clear		NOWT	Weight hhungs
use "Gloanel data analysid	, rename xPSUID PSUID	FWT	Integer weights in
or other one entropy	1	INDEWT	Integer weights hh
		DISTO	District ID contect
	. rename xHHID HHID2005	0.0.11	
		Properties	
	rename xRKSPLITID RKSPLITID2005		
		 Variables 	
	. sort STATEID DISTID FSUID EMID2005 MESFLITID2005	Name	
		Label	
	. save nn_incai	r)pe	
	real management and party	Value Label	
	. clear	Notes	
		4 Data	
	. use "G:\panel data analysis\IRD52_EH.dta"	Filename	
	(india muman bevelopment survey-11 (inus-11), 2011-12, mousenoid)	Label	
		 Notes 	
		Variables	758
		Observations	
	Command	Size	109.74M
	SOIT STATEID FSUID HMID HMISFLITID	Memory	160M
		Sorted by	STATEID DISTID
tata15			GR NUM



So, we will open the household two. Alright. This one, so this has been opened. It is in front of you, we will sort the important variables, those common IDs. Alright. So, this has already been sort. So, this step for you have now guided to you so sort all those indicator or important identifier, unique variables. We also check, check the IHDS ID as well. Just a minute, we have already sorted it out.

(Refer Slide Time: 26:10)



We can check their IHDS also, whether they are uniquely identified or not.

(Refer Slide Time: 26:13)

leview T 3 ×	rename xSTATEID STATEID	Variables	T Q ·
City community have		A Discourse	
Intercomments nere		The same	ves neve
Command Jrc	. rename adistid distid	name	Label
use unpaner data analysis)		SURVET	MOS-1 (2005) of L.
rename «STATEID STATEID	. rename xPSUID PSUID	OKTO	HQU state code
rename (DIKTED DIKTED		0010	HQ0 PS11 village/
rename x00110 00110		HHD	HOO Hourshold I
rename xi JUD VIJD2005	. ATIMIN ANALY MALEVAL	HINGO	100 Sold hoursh
rename villiS0 ITTD MUSD		IDESU	PSU id unique 6 di
sort STATEID DISTID PSUID	. rename xEMSPLITID EMSPLITID2005	IDHH	Household id uni
size hh ibds]		WT	Weights in IHDS-II
0 clear	. SOFT STATELD DISTID FSUID WHIDZOUS WHEFLITIDZOUS	NOWT	Weights hh->pop
1 use 'G\panel data analysis).	save hh ihds1	FWT	Integer weights in
sort STATEID PSUID HHID	file hh_ihds1.dta saved	INDEWT	Integer weights hh.
3 INIX STATELD PSUID HIND 459		015701	District ID convect.
4 sort STATEID DISTID PSUID	. clear		
5 isid STATEID DISTID PSUID	use (C.)mane) date analyzis/10002 00 dta:	Properties	
	(India Human Development Survey-11 (IHDS-II), 2011-12, Household)	<u> </u> •••	
		4 Variables	
	. sort STATEID PSUID HHID HHSPLITID	II Name	
		Label	HQU UISTICT CODE
	. Isid STATELD FAULD MELD MESPLITID variables CTATELD POLID MELD MECHINELD do not unionally identify the observations	type	101 101 101
	r(159);	Value Ishel	
		Notes	
	. sort STATEID DISTID PSUID HHID HHSPLITID	4 Data	
		Filename	DIDS2 HH dta
	. 1916 STATELD DISTLD PSOLD WALD WASPLITLD	Label	
		 Notes 	
		Variables	758
		Observations	
Da la	Lommand	Size	109.74M
		Memory	160M

Simply will change a two IHDS. So, variable. These are not uniquely identifying the observations. Alright, State ID then district ID, PSUID then household, ID then split ID, 1, 2, 3, 4, 4 were only mentioned earlier. 4 to be 5, So, that is why there are some error mark. But it has been corrected. After understanding their unique identifiers, we have also saved and we have also sorted those file we will open the linking file, we will save it then save and replace then we will find the linking files.

(Refer Slide Time: 27:26)

Concast Internet (Conception) Conception	mund her v M v M v M v M v M v M v M v M	w Tax	. rename xXHID HHID2005	• V	/ariables	T 9
Commad Comma	<pre>M</pre>	iter commands here			Kiter variable	
with Compared the support - with STATED DISTOR ProDUCE MEDICING MEDICIN	Add an angle. Add and angle. Add angle.	Command III	rename #MKSPLITID #MSSPLITID2005		Name	Label
 Her STATE DISTRIPTION RELATIONS REFERENCES Her STATE DISTRIPTION REFERENCES HER S		use 'G/papel data analysis(STATEID	STATED
name Add 100 A	ANNO LAND AND AND AND AND AND AND AND AND AND	rename 1 x1	. sort STATEID DISTID PSUID HHID2005 HHSPLITID2005		0/5110	015110
<pre>maxes dbb Dbb Dbb Dbb Dbb Dbb Dbb Dbb Dbb Dbb</pre>	Set Definition of the set of the	rename xSTATEID STATEID	and by (bda)		PSUID	PSUID
<pre>marker #000 PADS marker #000 PADS m</pre>	MADE PAGE	rename xDISTID DISTID	file hh ihds1.dta saved		HHID	HHID
<pre>inter workstop w</pre>	•••••• •••••	rename xPSUID PSUID			HHSPUTTD	HHSPUTID
amer avoid 100 0000, wold and 2 hand 200 mole, wold the same breek parts 11 1000 1000 mole, wold and 2 hand 2 h	SeedEll Double	rename xHHID HHID2005	. clear		HHID2005	HHID2005
addition DDD PROD. were wall der were wall der were wall der were wall der were wall der were wall der wertsette Froz beite Bester strichte Froz beite Bester der wer statte Dollow Bester er der wer statte Dollow Bester der wer statte Dollow Bester der der wer schemen der wer der der der der der der der d	NED DOED PARTY IN THE SAME SERVICES AND SERV	rename xHHSPLITID HHSPL	and the second data and a second states and the second sec		HHSPUTTD20.	. HHSPLITID2005
use Nonit First STATUS FRUD RELATION First STATUS FRUD RELATION use Vacant status relation fruits List STATUS FRUD RELATION First STATUS FRUD RELATION use Nonit Status relation fruits List STATUS FRUD RELATION First STATUS FRUD RELATION use Nonit Status relation fruits List STATUS FRUD RELATION RELATION First Status relation fruits use Nonit Status relation fruits List STATUS FRUD RELATION RELATION First Status relation fruits use Name List STATUS DISTID FRUD REL REFERITOR List STATUS DISTID FRUD RELATION use Name List STATUS DISTID FRUD RELATION List STATUS DISTID FRUD RELATION use Name List STATUS DISTID FRUD RELATION List STATUS DISTID FRUD RELATION use Name List STATUS DISTID FRUD RELATION List STATUS DISTID FRUD RELATION use Name List STATUS DISTID FRUD RELATION List STATUS DISTID FRUD RELATION use NAME List STATUS DISTID FRUD RELATION List STATUS DISTID FRUD RELATION use NAME List STATUS FRUD DISTID FRUD RELATION List Internation use NAME List STATUS FRUD DISTID FRUD RELATION List Internation use STATUS FRUD DISTID FRUD RELATION List Internation Nonit use STATUS FRUD DISTID FRUD RELATION List Internation Nonit use STATUS FRUD PRUD DISTID FRUD RELATION Nonit Nonit	MAIL Pet STREED FUED RED REFAILING Seed STREED HID PADD MODE. HID STREED HID PADD MODE. HID PADD MODE.	sort STATEID DISTID PSUID	(India Human Development Survey-II (IHDS-II), 2011-12, Household)		HHWAVES	HHWAVES
Out or Vignation Support	and dat Angle. Hot Status Parks and Status Parks Base Status BAD Man Donoto Parks and dat Angle. Hot Status Data Data Base Status and dat Angle. Hot Status Data Data Data Base Status and data Angle. Hot Status Data Data Base Status angle data Bas	save hh_ihds1				
aid Topendia Aurophi, and Table Physics Mail Mail Topendia Aurophi, and Table Physics Mail Mail Topendia Aurophi, and Table Physics Mail Mail Topendia Aurophi, and Table Physics Physics Physics Mail Mail Topendia Aurophi, and Table Physics Physics Physics Mail Mail Topendia Aurophi, and Table Physics Phys	and dia ward. How NGB Med. HOW NGB MED. HIGH STRATED DISTO FAUD AND AND AND AND AND AND AND AND AND AN	clear	. sort STATEID PSUID KHID KHSPLITID			
and MARD NAB WOOD. and MA	NUB_PRODUCTION Intelligence of the set of the se	use 'G\panel data analysis\				
Add/Add/DAR/MARK FICS/P FICS/P FICS/P FICS/P and SALED DAR/DAR/MARK FICS/P FICS/P FICS/P FICS/P and SALED DAR/DAR/DAR/DAR FICS/P FICS/P FICS/P FICS/P and SALED DAR/DAR/DAR/DAR/DAR/DAR/DAR/DAR/DAR/DAR/	100 PAUD 0400 PAUD. F1(19): 110 PAUD 0400	sort STATEID PSUID HHID	variables STATEID PSUID HHID HHSPLITID do not uniquely identify the observations			
	HOLDAD PALL AND AND ADD PALL AND AND ADD PALL AND AND ADD PALL AND AND ADD PALL AND AND ADD PALL AND AND ADD PALL AND ADD ADD ADD ADD ADD AND AND ADD ADD	ISId STATEID PSUID HHID 459	r (459) :			
ang HARUADUM MAD and TRATED COTTO PUTD MED MERITID ang HARUADUM MAD and TRATED COTTO PUTD MED MERITID ang HARUADUM MAD AND AND MEDIAND MERITID ang HARUADUM AND AND AND MERITID ang HARUADUM AND	Million State and threft Distrib FirdD SKID KHIPITD and threft Distrib FirdD SKID KHIPITD 	SOR STATED USTIC PSUID.		P	roperties	4
الفلة 1940 00000. الفلة 1940 00000. الفلة 1940 00000. الملح 1940 000000. الملح 1940 0000000000000000000000000000000000	- 1416 STARTIS DISTID FMUTD BKID KKEPLITID VALUG '0' (pasel data makjris/1005)_88.48* * * *********************************		. sort STATEID DISTID PSUID HHID HHEPLITID		ê • •	
aver till de () space/discupped. aver till de () space/discupped. aver/ikkliD PNUD DODD. clear is de () space/discupped. is de () space/discupped. - clear - clear is de () space/discupped. is de () space/discupped. - vam *0 () space/discupped. - clear vam *0 () space/discupped. is de () space/discupped. - vam *0 () space/discupped. = set statting PNUD DISTO PNUD BRID REPORT P De () space/discupped. is der () space/discupped. - vam *0 () space/discupped. = set statting PNUD DISTO PNUD BRID REPORT is der () space/discupped. is der () space/discupped. - vam *0 () space/discupped. - vam *0 () space/discupped. - vam *0 () space/discupped. is der () space/discupped. - vam *0 () space/discupped. - vam *0 () space/discupped. - vam *0 () space/discupped. - vam *0 () space/discupped. - vam *0 () space/discupped.		clear	1916 STATEID DISTID FSUID #HID #HID HESPLITID		 Variables 	
ANN TAILD PHILD DOIDD. - NAW, 1923ac (Lot ULAL) 11 (G () pask) data analysis/1052; 98.05 a saved 1 Norm Norm class: - use "Q () pask) data analysis/1052; 98.05 a saved 1 Norm Norm . use "Q () pask) data analysis/1052; 98.05 a saved 1 Norm Norm 1 Norm . use "Q () pask) data analysis/1052; 98.05 a saved 1 Norm Norm 1 Norm . use "Q () pask) data analysis/1050; 01070; 98.05 a saved 1 Norm Norm 1 Norm . use "Q () pask) data analysis/1050; 01070; 98.05 a saved 1 Norm Norm Norm . commod 1 Norm 1 Norm Norm 1 Norm	ID PHUD DODD.	use "Glipanel data analysis"			Name	
Tile Gigandi data analysis/18072 %.dta avved 1000 - claar - var "C \pead data analysis/110815 dat" - set STREED FUED DIFFED FUED SUBSE SELECTED - Var - Command - Command	It is () () () () () () () () () () () () ()	sort STATEID PSUID DISTID	save, replace			
. clasz 1998 Gold - 1998 G	- clast - clast - use *0 'gand data analysis'licabh das* - seut STATECO FAUD DIFEO FAUD		file G:\panel data analysis\IHDS2_HH.dta saved		type	
use "G \parts (data analysis)(licit)h das" sect STATED PRUTO DISTED PRUTO MED MERINITED Command 1	سورج ۱۱ SINIED PHUE MEE MARFIELD wing "O (panel data analysis/1925) (M. da" سورج ۱۱ SINIED سورج ۱۱ SINIED PHUE DITTE PHUE MEE MARFIELD سورج ۱۱ SINIED سورج ۱۱ SINIED PHUE DITTE PHUE MEE MARFIELD سورج سورج ۱۱ SINIED سورج سورج سورج سورج سورج سورج سورج سورج سورج سورج سورج سورج سورج سورج سورج سورج سورج سورج سورج		, clear			
umm 0 (pass) data malgravitation). Mar • Des set Statuto • Des • Locar . • Mont • Locar . • Mont • Command	um 0 (pase) data malyrariticable. fer" 4 000 set STATED FAUD DISTO FAUD DISTO FAUD DISTO MALO DISTO FAUD					
sect STATED FUID DISED FUID AND WESKITED AND AND AND AND AND AND AND AND AND AN	set STREED FEUD DISTD FOUD MED MERFLETD Main Fill Command Command Command Marger 11 STREED FEUD DISTD FOUD MED MERFLETD Street Total Command		. use "G:\panel data analysis\linkhh.dta"		- Data	
Lee Lee Some Command Comma	Lade • Communid • Communid • ************************************		ANT PRET BUTT DITAT DETTA BUTT BUTT BUTT			hnkhhdta
	. " Non - Vorden 8 Vorden 9 V				Label	
Command Command	Command 4132 Command 4132 Sor 17046 Sor 17046 Meneys 11 STATID DISTD FAUD HEIPLITTD water 'G-\pacel data analysis\TDDS_ME data' Meney 423 Sor 37046 Meney 423 Sord 14				Notes	
Command 4210 See 1104	Commund 8 0mmund 8 0mmund 5/m 10.94% merge 11 STATID DISTD MUD MUD MUDIUM (%) (mmmund 6/m 0mmund 4/m 0mmund 4/m				Variables	8
Size 3304	Ser 10544 Menege 1:1 STATED DISTED FAUED EMSFLITED wing "G'genel data analysis/1855]ems dta" Menoy 4A4 Sonety StATED PAG		Command		Observations	42,152
merce 1:1 STATEDD DISTUD FOUD HAID HASPLITID USING "G:\Danel Gata analysis\IHDS2 HH.Gta"	Memory GAM Sorted by STATED PSU		merge 1:1 STATEID DISTID FSUID HHID HHSFLITID using "G:\panel data analysis\IHDS2 HH dta"			370.48K
Menoy GM						

Linking files are here. Let me just open, It is here. Then we will go for understanding the linking files. It is linked HHID. Alright. So, the linking file is already given in IHDS website for our clarification. I just wanted to give some information that it gives the unique identifiers the state ID, district ID, PSUID. , interestingly, it gives information about HHID.

And split ID of that without the name of 2012, but only given HHID. This means it is of 2012 household ID and split ID, households split ID that mean this is of 2012 and these are these last two are of household ID of 2005 and household split ID of 2005. So, while we merge the two dataset we have to take very carefully about which linking files we are supposed to connect.

(Refer Slide Time: 28:47)

	isid STATEID DISTID PSUID HHID HHSPLITID
	Step 8: save the sorted file
	save, replace
	Step 9: open linking file
	use linkhh.dta, clear
	Step 10: sort linking IDs in linking file
	sort STATEID DISTID PSUID HHID HHSPLITID
	Step 11: merge IHDS-II household file to household linking file
	merge 1:1 STATEID DISTID PSUID HHID HHSPLITID using
	"IHDS2_HH", gen (_mergeR2link)
0	Swayan 6

Our important linking files for merging two datasets like linking file with IHDS-2. So, we have to sort those 5 ID. So, as per 2011-12 data, these three, four, five. Alright. So, we have already sorted out. We can also check the IHDS ID as well. But since it has already been defined so no need to check. We will simply merge, merge it here,

(Refer Slide Time: 29:30)

	Immediate intermediate intermedinte intermediate intermediate intermediate int	In the second seco
	Control Firster SSFULTID SSFULTID2005 String Control Firster SSFULTID SSFULTID2005 String Control String Control String Control String Contro String Control <th>Constant of the second se</th>	Constant of the second se
	Provide data analysis Provide data	35.000 55.000 55.000 55.000 55.000 55.000 55.000 55.000 1000000 1000000 1000000 1000000 10000000 1000000 10000000 1000000 10000000 1000000 10000000 1000000 1000000000 1000000 10000000000000 1000000 1000000000000000000000000000000000000
	<pre>set 1 ATTO D DITTO FROD BADDONS BARFITTD D THE SET AND D DITTO FROD BADDONS BAFFITTD D THE SET AND D DITTO FROD BADDONS BADFATTD D THE SET AND D DITTO FROD BADDONS BADFATTD D THE SET AND D DITTO FROD BADDONS BADFATTD D THE SET AND D DITTO FROD BADDONS BADFATTD D THE SET AND D DITTO FROD BADDONS BADFATTD D THE SET AND D DITTO FROD BADDONS BADFATTD D THE SET AND D DITTO FROD BADDONS BADFATTD D THE SET AND D DITTO FROD BADDONS BADFATTD D THE SET AND D DITTO FROD BADDONS BADFATTD D THE SET AND D DITTO FROD BADDONS BADFATTD D THE SET AND D DITTO FROD BADDONS BADFATTD D THE SET AND D DITTO FROD BADFATTD D THE SET AND D DITTO FROD BADFATTD D THE SET AND D DITTO FROD BADFATTO D THE SET AND D DITTO FROD BADFA</pre>	CODIO CONTO HIGO DE LA
<pre>remove control to the second control to</pre>	<pre>rever CMU DDUTU rever CMU DDUTU rever CMU DDUTU rever CMU DDUTU rever CMU DDUTUU rever CMU DDUTUU rever CMU DDUTUUU rever CMU DDUTUUU rever CMU DDUTUUUU rever CMU DDUTUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU</pre>	System Paulo Paulo Herodunic Herodunic Herodunic Heroduni
<pre>result of the subset is nevel is an initial time is a subset is nevel is an initial time is a subset is nevel is an initial time is a subset is nevel is an initial time is a subset is nevel is an initial time is a subset is nevel is an initial time is a subset is nevel is nevel is a subset is nevel is neve</pre>	<pre>rever det 00000 rever rever det 0000 rever re</pre>	HIND HIND HIND HIGHLING HINDLENG HINDL
<pre>served.op.dop.op. served.</pre>	<pre>stars.wd:U00 H000 waren.wd:U00 H000 waren.wd:U00 H000 waren.wd:U00 H000 waren.wd:U00 H000 waren.wd:U00 H000 waren.wd:U00 H000 wd:U000 H000 wd:U0000 wd:U000 H000 wd:U0000 H000 wd:U000</pre>	HIGUID HIGUID HIGUDS
<pre>start with Units Decision is an addition to be addition to be</pre>	<pre>interest #04.000 #04.000 with MADD 000 PAPOL WIth MADD 000 PA</pre>	 Headdard and the second second
<pre>set of the set of</pre>	<pre>state of the state of the</pre>	HIGHTOD, HIG
<pre>withing controls. withing controls. withing</pre>	And District Processor India manas Development Survey-11 (DBS-11), 201-12, Bouehold) Get - set STATES PROD BUS MARIESS With Mind District Processor - set STATES Processor States Processor - set STATES Processor With Mind District Processor - set STATES Processor States Processor - set STATES Processor With Mind District Processor - set STATES Processor States Processor - set STATES Processor With Mind District Processor - set STATES Processor States Processor - set States Processor	Hereau Galaxies
<pre>structure s</pre>	Are Mynoli of Gued at A waytor. With Million Mullion With Herits Freide Mark Market Hill With Million Mullion With Herits Herits Mark Market Hill With Million Mullion With Herits Herits Mark Market Hill With Million Mullion With Herits Herits Mark Mark Hillion With Million Mullion With Herits Herits Mark Mark Hillion With Million Mullion With Herits With Million Mullion Mulli	Nopries Norme
<pre>view view view view view view view view</pre>	<pre>view dia natyri. view di</pre>	Popries • Virial- •
<pre>Number of the second of t</pre>	<pre>multiple provide set SIMID Provide set SIMID</pre>	Noperies Image: image ima
<pre>Willingsonderde Willings</pre>	WILMID DATA DUMBLE, edited in the second	Properties Visionia Late webutt Late webutt Late webutt Some
<pre>struture control. set spins set spin set spin set</pre>	ar SUND DOTOP PUD. ar roku: ar VSUND DOTOP PUD. ar roku: ar VSUND DOTOP PUD. ar VSUND PUD. DOTOP (is of VSUND data analysis)1002 BM.dts aread ar VSUND PUD. DOTOP PUD. ar VSUND P	Poperies
<pre>widthings provides and provide widthings widthings</pre>	will MANDOR PUPOL wie Taking Diffic Fault Bails Bails Bails Bails Bails Wigen da work. wissen of the Signed data analysishiddly did a sevel - clear - c	A state of the second sec
<pre>set upget set up</pre>	And region or Support dia serversit. art SUITO PROD DOTO:	Product Margin Mar
<pre>inf upwel at winty:</pre>	der wigend die anoppin. erst MANDE PARE DOTTO: Wie 1990 BUD BAUR KAFFLITED erst MANDE PARE DOTTO: erst Filme Stylpend date analysis/10002_00.dts averd - clear - erst Simtle PARE DOTTO BAUR KAFFLITED erst Simtle PARE DOTTO BAUR KAFFLITED erst Simtle PARE DOTTO BAUR KAFFLITED ASS CONVENDENT CON	American Sectors and the sector of the secto
er uppen en envel. er uppen en envel. er en envel. er en envel. er envel.	Ar "Gynend dia anarysis. Hirds Type To Bala Hirds Type To Hirds Hirds Hirds Type To Hirds Hi	Lake rectange of the second seco
ett status pado dobb. int i singen da sanjeri lidad da sanjeri lidad da si int i singen da sanjeri lidad da sanjeri lidad da int i singen da sanjeri lidad da si int i singen da sanjeri lidad da sanjeri	ALL SALED PAOL COID.	I get and Mag Voca Lane Mag Voca Lane Mag Voca Lane Mag Anno Anno Voca Voca Voca Voca Voca Voca Voca Voc
<pre> first if if if</pre>	- cierz	format with a Maly format under format
<pre> i iiii i iii i iii i iiii i iiiiii i iiiiii</pre>	All Die Gegen den angegenhenen All Die Gegen den angegen den	Vision Labor Vision
<pre> test Solution for the solution</pre>	All 12 Connect data analysis/1002 Without the All 12 Connect data analy	Kon K
<pre> set SANDED FANDE GUIDE FANDE MARKITED inst SANDED F</pre>		Anner Anner Anner Vanden Nore Vanden
<pre>inf Shalls Field Shalls Shall S</pre>	All 15 Constant and provide the set of the s	Luter Products II Communities III Source Datas Bernary Sub Lateron - 1 0 (1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
A set of the set of	All All All All All All All 10 (varead an anyweighterhold) (a) So (a) (b) (b) (b) (b) (b) (b) (b) (b) (b) (b	Mon Yone Convertors
All Control All Control	Command Rege 1:1 a5 C C C C C C C C C C C C C	Vanisten 8 Observations C1122 Menory 6MM Sonte by SIATE CAR • 2 © 4 to № 42 101
Ale and a second and a second and a second at a s	ACCESS Connect rege 1.1 ACCESS Connect Conne	Conversion CLS2 Sector y States Sector y States CCC CCC CCC CCC CCC CCC CCC
<pre>sty is if if the set is if if the set is if if the set is if the se</pre>	AS AS AS AS AS AS AS AS AS AS	Size Disast Menony GM Sorted by STATED CAP 11/1 11/1
All and a second and a second at a second	Al 111 Connect Ans many extent of Al 111 Connect Ans many extent on Al 111 Connect Ans many extent on Al 111 Connect Ans and part of the analysis (1000 1 M 401) M → A + C + M → C + Q + Q + Q + Q + Q + Q + Q + Q + Q +	Menoy 6M Soted by Static + 1 • 1 • 1 1 1 1 1
30 Set	AS AND IS Conception and any encoder Set Disc Capability Entries Die Worken Hop Set Sin & J - M - Set Sin Die Origination Strangent/SEGS_M. des- Mark Set Sin Set Sin Die Origination Strangent/SEGS_M. des- Mark Set Sin	- 1 0 1 10 10 100
20 20 <t< td=""><td>AS AS 151 - Compared Assembled interments Bet 151 - Compared Assembled interments Bet 151 - Compared Assembled interments Bet 151 - Compared Assembled interments W = 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1</td><td>- 2 0 € to h 10 10</td></t<>	AS AS 151 - Compared Assembled interments Bet 151 - Compared Assembled interments Bet 151 - Compared Assembled interments Bet 151 - Compared Assembled interments W = 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1	- 2 0 € to h 10 10
<pre>13 U S . 4 S . S . S . O .</pre>	B U S · <u>u</u> - <u>S</u> · · <u>S</u> ·	
w Y # Z * Viable 12 File (Note: 12 Fi	w Y 2 K we "Channel data malgiral[DD2] M dtt" dreinsmach we O Greinsmach we O set STMTLE PROF DB DE MERTID	
An contractive Image: Status and an experiment of status and and experiment of status and experiment of	terconnaditer (India nuban vereighent auregrit (India 11), ettiltz, neusenoid) (onnad (K), soit STATEID FSUID XHID XHID XHIFTIDD	 Variables
Conversion **** **** **** **** ***** Set 10 STATUS FUED WEDD STATUS THE SET STATUS THE	Command rc . sort STATELD PSVID WHID WHID WHISPLITED	 Filter variables here
<pre>ind Superdia standyrd, marker if MIDE STATUS STATUS and STATUS and STATUS and STATUS marker MIDE STATUS STATUS STATUS AND STATUS marker MIDE STATUS STATUS STATUS STATUS STATUS STATUS marker MIDE STATUS STATUS STATUS STATUS STATUS STATUS marker MIDE STATUS STATUS STATUS STATUS STATUS STATUS STATUS STATUS marker MIDE STATUS STATUS</pre>		Name Label
<pre>maxes ** ** ** ** ** ** ** ** ** ** ** ** **</pre>	use "Gripanei data analycis"	STATEID STATEID
<pre>remer clifts 1000 memor clifts 1000 remer c</pre>	reame's' . Isid STATED POULD WID WESPLITED	DISTID DISTID
<pre>state dBDD 2000 mater ethol 200 mater et</pre>	resame xSTAHED STAHED ELECTRON CONTRACTOR OF THE CONTRACT CONTRACT CONTRACTOR CONTRACTOR CONTRACTOR	PSUID PSUID
<pre>marker.dH.QU.B.W.D.W.D</pre>	x0ISTID 0ISTID	A HHID HHID
<pre>mark refu](D) Multiple state refu](D) Multiple st</pre>	reame VSUD SUD . sort STATEID DISTID FSUD KKID KKSPLITID	HHSPUTTD HHSPUTTD
<pre>set static provide the set of the set o</pre>	NAME WHICH HID/NDS	HHU2005 HHU2005
and mAndl are mAndl are many many many many many many many many	same Annoulli Angu.	HEWAVES HEWAVES
Operation - Class 0000 HPU descent	. save, replace Swenhulds1 file G:\mamel data analysis\THD02 HK.dta saved	SURVEY IHOS-1 (200
ar "Gyaned a supplication of the second of t	lear in the original and analyzes incor and	IDPSU PSU id unig
et MAIR DOWNED. et MILEO NOW HELE et MILEO NOW HE	ise "Glyanei data analysis"	IDHH Household
WILKING DOOD, OLE, AND STATUS MUST AND	ext STATEID PSUID HHID	W1 Weights in I
et VIAID 0000 704/DL et VIAID 0000 704/DL et very all ser very all set to total 000 set of the accompacts willing data's values) (set very alle SETED var byte, now lat to accompact willing data's values) (set very alle SETED var byte, now lat to accompact willing data's values) (set very alle SETED var byte, now lat to accompact willing data's values) (set very alle SETED var byte, now lat to accompact willing data's values) (set very alle SETED var byte, now lat to accompact willing data's values) (set very alle SETED var byte, now lat to accompact willing data's values) (set very alle SETED var byte, now lat to accompact willing data's values) (set very alle SETED var byte, now lat to accompact willing data's values) (set very alle SETED var byte, now lat to accompact willing data's values) (set very alle SETED var byte, now lat to accompact willing data's values) (set very alle SETED var byte, now lat to accompact values) set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, now lat to accompact values) (set very alle SETED var byte, n	sd STATED PSUD HHD. 459 . use "G:\panel data analysis\linkhh.dta"	INDWT Weights hh
MarkINDORDONUEL. wergen 11 strattig Strattig sample, mov list to accommodate uning data's values) Mergen it sample. Mergen 11 strattig Strattig sample, mov list to accommodate uning data's values) Mergen 11 strattig Strattig sample. Mergen 11 strattig Str	ert STATEID OSTBO PSUD sost STATEID FSUID DISTID FSUID KKID KKEPLITID	Properties
instruction	sid STATED OISTID PSUID	
are consistent analysis. disponentiate analysis. extraction products analysis. extraction products of the construction o	ant. replace marge 1:1 STATED DISTED FUED REID RESPLITED using "G'(panel data analysis)IRD52_EM.dta"	 Variables
an experimentation (note: viriable RETITURATION reper 113ARTO DOND). Reper 113ARTO DOND. Reper 113ARTO DOND.	(note: variable DISTID was byte, now int to accommodate using data's values)	Name STATED
Instr: writikie MRITITIO us byte, nov ist to accomposite using data's values) instr: instr	(note: variable PSUD was byte, now int to accommodate using data's values)	Label STATED
Interface Interface <t< td=""><td>menge 1: STATED OISTD_ (note: variable RESPLITED was byte, now int to accommodate using data's values)</td><td>Type int</td></t<>	menge 1: STATED OISTD_ (note: variable RESPLITED was byte, now int to accommodate using data's values)	Type int
not satched 0 natched 47,332 (netpres))	Result # of obs.	Volum Labor
net statued 0 Part natched 42,152 (metgrewit) Family General Genera Genera Genera <td></td> <td>Notes</td>		Notes
matches U_112 (Metgen-1) Instruct <	not matched 0	4 Data
ب الملاح مرتب مرتب مرتب مرتب مرتب مرتب مرتب مرتب	natched 42,152 (_merge==3)	Filename linkthidta
		Label
Command 100 Command 100 Sam 10000 Wenny 1000 1000		 Notes
Command 4 (see 1930) 5.00 1930 Menny 104		Variables 762
Size 19900 Menory 1604	Command	Observations 42.152
- Menoy 16M		
		Size 109.90M

So, it is basically 1 to 1 merging and 3. So, this is 5 IDs using our linking file. So, using our IHDS-2 here. You can close. So, this we have already sorted out all its unique identifier, alright. It is already opened. So, it is here then once again we will open. Alright. So, IHDS-2 we have used now there is a merging, we have gone through the merging. The merging has already been completed at this moment.

You can mark, which we have guided from the beginning, that 42152 are the total observation in IHDS-2 and the linking file also the entire information were also available for and that to we followed one to one merging. So, that is very clearly understood and our merging has been

successful because of this linking file with IHDS-2. So, what we will do, we have a combined IHDS-2 with linking file information has already been included in this file. So, what we will do will then generate a merge file.

(Refer Slide Time: 31:22)

isid STATEID DISTID PSUID HHID HHSPLITID	
Step 8: save the sorted file	
save, replace	
Step 9: open linking file	
use linkhh.dta, clear	
Step 10: sort linking IDs in linking file	
sort STATEID DISTID PSUID HHID HHSPLITID	
Step 11: merge IHDS-II household file to household linking file	
merge 1:1 STATEID DISTID PSUID HHID HHSPLITID usi	ng
"IHDS2_HH", gen (_mergeR2link)	
🍥 swojam 🔮	18

So far since we are trying to link and merge the round 2 dataset, we can generate file with this name. So, already generated.

(Refer Slide Time: 31:40)

riew 🕇 🕯	. use "C:\panel data analysis\18052_88.dta"	 Variables 	τ 0
Filter commands here	(India Human Development Survey-II (IHDS-II), 2011-12, Household)	+ Filter valiab	
Command rc	SOIT STATEID PSUID WHID WHIPHITID	Name	Label
use 'G\panel data analysis_		STATEID	STATEID
rename * x*	. isid STATEID PSUID HHID HHSPLITID	0510	DISTID
rename xSTATEID STATEID	Variables STATEID PSUID HHID HHESPLITID do not uniquely identify the observations	PSUID	PSUID
rename xDISTID DISTID	1 (403) 1	 HHID 	HHID
rename xPSUID PSUID	sort STATEID DISTID PSUID HHID HHSPLITID	HHSPLITID	ннорато
rename xHHID HHID2005		HHID2005	HHID2005
rename xHHSPUTTD HHSPL.	. 1914 STATEID DISTID PSUID HHID HHSPLITID	HKSPUTID20	HHSPUTID2005
sert STATED DISTID PSUID.		HEWAVES	HHWAVES
save hh inds1	. save, replace	SURVEY	IHDS-1 (2005) or L
clear	tile o:/panel data amalysis/ikust_kk.dta saved	IDPSU	PSU id unique 6 di
use "Ghoanel data analysish	, clear	IDHH.	Household id, uni-
sort STATEID PSUID HHID		wf	Weights in IHOS-II
ISI STATED PSUID HHID 459	. use "G:\panel data analysis\linkhh.dta"	INOW?	Weights hh->pop
sort STATEID DISTID PSUID.			
isid STATEID DISTID PSUID	. SOIT STATEID PSUID DISTID FSUID MMID MMIFLITID	Properties	
cave replace	manuel 1:1 STATEID DISTID FUID MAID MAIPLITID using "G'hanel data analysis\IMD52 MA dta"	<u> </u>	
clear	(note: variable STATEID was byte, now int to accommodate using data's values)	* Variables	
use "Ghoanel data analysish-	(note: variable DISTID was byte, now int to accommodate using data's values)	Name	
sert STATED PSUID DISTID	(note: variable PSUID was byte, now int to accommodate using data's values)	Label	
menne 1:1 STATEID DISTID	(note: variable RESPLITID was byte, now int to accommodate using data's values)	II Type	
	Besult & of obs	Format	58.09
		Value label	
	not matched 0	Notes	
	matched 42,152 (_merge==3)	* Data	
		riename Label	instructu
		La Caber	
		• Notes	
		Chanatana	
	Command	Costinations	
		Memory	160M

I can show you that is already there at the end. So, this can be renamed with the merge. Since merge either merge is to right link? So, we have created a variable with this particular name. Alright. So, we will come back to this further clarification.

(Refer Slide Time: 32:31)

Re no ma	ult : matched sched	# of obs. 0 42,152 (_merg	te22link==3)	Here, you can so observations have matched because ID identifying in both and has equal n observation	ee all the completely s are uniquely the dataset umber of ins.
Step	12 : sort the	id variable fo	or linking IH	DS-1 file and sa	ive this
link a	nd wave 2 fi	le with differe	ent name.		
sort S	TATEID DIST	ID PSUID HHI	D2005 HHSF	PLITID2005	
save	hds2_link				
Step	13 : open IH[DS-I househol	d file		
use h	h_ihds1.dta	, clear			
🎯 - swayam	9				19

So, this is what we have guided. And here you can see all the observations that have already been completely merged because IDs are uniquely identifying both the datasets and has equal number of observations. Coming to the next step, that sort the ID variables of IHDS-1 and the IHDS-2 and the linking file, we have already merged. We have a combined dataset of that.

We need to get the IHDS-1 into it. So, IHDS-1 we already saved by sorting their unique identifiers. We will save that then we have two file with different name. There will also save it accordingly. What is suggested here that you need to sort identifier of IHDS-1, that we have already done it.

So, we saved it with this name as well. But this is the file we have created basically the merge file we have created. But we will try to merge with IHDS-1. IHDS-1 contains the unique file and information like this and this. Since this file has not been sorted with this information. First duty is to sort with these names, these variables, then only we can able to merge correctly. So, let us sort these five indicators.

(Refer Slide Time: 34:11)



So, sort just a minute. It is here, so 1, 2, 3, then 5. So, we have sorted out. So, this we are naming it, save with IHDS-2 with link IHDS-2 and link file has already been saved. So, what is the next important aspect we will open the IHDS-1 household file that those IDs we have already sorted in IHDS-1. So, we need not sort further.

(Refer Slide Time: 35:00)



We will open that first then we will merge with the IHDS-2 link file. So, let me first clear this out. So, we will clear now will open IHDS-1 Household file.

(Refer Slide Time: 35:19)



So, IHDS-1 and that we have already sorted has been opened. Now, what will do? will merge.

(Refer Slide Time: 35:43)



So, we have used this file, isn't it? And I think we have already sorted it.

(Refer Slide Time: 35:49)



All those indicators. So, we are supposed to go for merging. So, we will merge it with this information, we will straight away take this command, straight away copied. Why copied because we will carry with the same name generate the way we defined earlier will go by that only. And this is here.

(Refer Slide Time: 36:20)

leview T 3		Variables	τ
Filter commands here	. save, replace	+ Filter variabl	
f Command In	file 6:\panel data amaiysis\imid2_mm.dta saved	Name	Label
use "Gloanel data analysis).	. clear	xCASEID	HH ID: 11 byte stri
rename * x*		STATED	State code
rename xSTATEID STATEID	. use "G:\panel data analysis\linkhh.dta"	DISTID	District code
rename xDISTID DISTID		PSUID	PSU village/neigh
rename xPSUID PSUID	. Pole similio foolo sistio foolo milo mortilio	HHID2005	Household ID
rename xHHID HHID2005	. merge 1:1 STATEID DISTID FSUID HEID HESPLITID using "G:\panel data analysis\IEDS2_EE.dta"	HHSPLITID20	. Split household ID
rename xHHSPLITID HHSPL	(note: variable STATEID was byte, now int to accommodate using data's values)	XDHH	HH ID 9-digit unig
sort STATEID DISTID PSUID.	(note: variable DISTID was byte, now int to accommodate using data's values)	xIDPSU	PSUID 6-dipit uni.
save hh inds]	(note: variable PSUID was byte, now int to accommodate using data's values)	xSTATEID2	State codes, colla
0 clear	(note: variable nmortillo was byte, now int to accommodate using data 5 values)	XDISTNAME	District codes with
1 use "G\panel data analysis\	Result # of obs.	xDIST01	H1sp: District ID C
2 sort STATEID PSUID HHID		XURBAN	Census 2001 villa.
3 Isid STATEID PSUID HHID 459	not matched 0	xMETRO6	Largest 6 metro ar
4 sort STATEID DISTID PSUID	matched 42,152 (_merge==3)		
5 isid STATEID DISTID PSUID		Properties	•
6 save, replace	. rename merge mergeR21ink		
7 clear		* Variables	
8 use "G\panel data analysis\	. sort STATEID DISTID FSUID HHID2005 HHSPLITID2005	Name	
9 sort STATEID PSUID DISTID	and the first state	Laber	conterent values :
0 merge 1:1 STATEID DISTID	. seve inds_line file (bds2 link dte saved	Type	
1 rename (merge (mergeR2)).		E Volution	
2 sort STATEID DISTID PSUID	. clear	Notes	
3 save ihds2_link		4 Data	
4 clear	. use "G:\Stata15\hh_ihds1.dta"	Filename	Nh ihdslidta
5 use "G\Stata15\hh_ihds1.dt.	(India Numan Development Survey (INDS), 2005, Nousehold)	Label	
		 Notes 	
		Vanables	945
		Observations	41,554
	Command	Size	100.97M
	marge indez link file with INDS-I file	Memory	160M
	merge 1:m STATELD DISTLD FSULD ARLDZOUS ANSFLITUZOUS USINg "INDSZ_INK.dtä", gen (_mergeKIKZ)	Sorted by	STATED DISTID P

Image: Solution (Section	eview T 🛛 🗙	(label CO6B already defined)		Variable	5	τa
<pre> (ormed</pre>	Filter commands here	(label CO5B already defined)		× 10		
image: strain strains, image: s	Command	(label CONB already defined)		Nor	lahel a	
a de generation indet total alressy defined indet total alressy defined a ware dotto Dotto (late) WD (alressy defined indet total alressy defined a ware dotto Dotto (late) WD (alressy defined indet total alressy defined a ware dotto Dotto (late) WD (alressy defined indet total alressy defined a ware dotto Dotto (late) WD (alressy defined indet total alressy defined a ware dotto Dotto (late) WD (alressy defined indet total alressy defined a ware dotto Dotto (late) WD (alressy defined indet total alressy defined a ware dotto Dotto (late) WD (alressy defined indet total alressy defined a ware dotto Dotto (late) WD (alressy defined indet total alressy defined a work dotto Dotto Dotto (late) WD (alressy defined indet total alressy defined a work dotto Dotto Dotto (late) WD (alressy defined indet total alressy defined a work dotto	use 'Glassel data analogid	(label CO2B already defined)		-CAS		11 hotestei
marker if Milling State intel it medic intersity of files South of the state South of the state marker if Milling State (Lake intersity intersity of files) South of the state South of the state marker if Milling State (Lake intersity intersity of files) South of the state South of the state marker if Milling State (Lake intersity intersity of files) South of the state South of the state marker if Milling State (Lake intersity intersity of files) South of the state South of the state marker if Milling State (Lake intersity intersity of files) South of the state South of the state marker if Milling State (Lake intersity intersity of files) South of the state South of the state cere (Lake intersity intersity of files) South of the state South of the state South of the state cere (Lake intersity intersity of files) South of the state South of the state South of the state cere (Lake intersity intersity of files) South of the state South of the state South of the state und State (Lake intersity intersity intersit) Intersity intersity intersit <t< td=""><td>use ongeneroatalanerport.</td><td>(label CO1B already defined)</td><td></td><td>100</td><td>10 1010</td><td>in oyie son.</td></t<>	use ongeneroatalanerport.	(label CO1B already defined)		100	10 1010	in oyie son.
marker ddbilloshindi marker dd	rename . x.	(label MM4C already defined)		SIAI	TID STRIG CO	01
marare:dBDDSDD (Late: WC:Latersky define) HDDB (Late: WC:Late: WC:Latersky define) HDDB (Late: WC:Latersky define) HDDB (Late: WC:Latersky define) HDDB (Late: WC:Late: WC:La	rename ISTATELD STATELD	(label MM3C already defined)		DDI	D District (code
marar #300 PXUS (List WC Listedy defined) marar #300 PXUS (List WC Listedy defined) (List WC List	rename xDISTID DISTID	(label MM2C already defined)		PSUI) PSU: vill	age/neigh
marar mildel 100005 marar mildel 100005 marar mildel 10005 marar mildel 10005 mara	rename xPSUID PSUID	(label MMIC already defined)		нн	2005 Househo	old ID
maramarangeng magnation body. Lines in ket and generalized body and the MED Sequence of the MED Sequence o	rename xHHID HHID2005	(label RC4 already defined)		HHS	UTID20 Split hos	usehold ID
up 111110 1000 107 100 1 1111 111111111111111111111111111111	rename xHHSPLITID HHSPL	(label RC2 already defined)		xIDH	H HHID9	-digit uniq
ue ret_A_ALI ever ever ue volume ever	sort STATEID DISTID PSUID	(label Kilb already defined)		xIDP	JU PSUID F	6-digit.uni
oper ver (landet strate) constant constant <td>save hh_ihds1</td> <td>(label KCI already defined)</td> <td></td> <td>ATZx</td> <td>TEID2 State co</td> <td>ides, colla</td>	save hh_ihds1	(label KCI already defined)		ATZx	TEID2 State co	ides, colla
we Visiting Pointers exponent set STATUD 20170 Point set Status 20170 Point	clear	(label NF22 already defined)		x015	INAME District	codes with
se Station Proponeet. se Station Proponeet.	use 'Gloanel data analysis).	(label NF2 already defined)		105	(01 H1ser D	istrict ID C
add STATUS DOS DDD. add Status dos Advisorials add Status dos Advisorials add Status dos Advisorials add Status DOS DDD. add Status dos Advisorials add Status dos Advisorials Baser das Advisorials add Status DOS DDD. add Status dos Advisorials add Status dos Advisorials Baser das Advisorials add Status DOS DDD. add Status dos Advisorials add Status dos Advisorials Baser das Advisorials add Status DOS DDD. add Status dos Advisorials add Status dos Advisorials Baser das Advisorials add Status DOS DDD. add Status dos Advisorials add Status dos Advisorials Baser das Advisorials add Status DOS DDD. add Status dos Advisorials add Status dos Advisorials Baser das Advisorials add Status DOS DDD. add Status dos Advisorials add Status dos Advisorials Baser das Advisorials add Status DOS DDD. add Status dos Advisorials add Status dos Advisorials Baser das Advisorials add Status DOS DDD. add Status dos Advisorials add Status dos Advisorials Baser das Advisorials add Status DOS DDD. add Status dos Advisorials add Status dos Advisorials Baser das Advisorials add Status DOS DD. add Status dos Advisorials add Status dos Advisorials Baser das Advisorials add Status dos Advisorials add Status dos Advisorials <td>sout STATEID PSUID HHID</td> <td>(label NF1B already defined)</td> <td></td> <td>107</td> <td>un Cenur</td> <td>2001 villa</td>	sout STATEID PSUID HHID	(label NF1B already defined)		107	un Cenur	2001 villa
Markawa Markawa <t< td=""><td>HIN STATED BUILD HIND 450</td><td>(label NFIA already defined)</td><td></td><td>-440</td><td>1006 Langert</td><td>6 matrix at</td></t<>	HIN STATED BUILD HIND 450	(label NFIA already defined)		-440	1006 Langert	6 matrix at
And instruction Drowing Lisket 11154 alressly defined Progettes Used statution Drowing Lisket 11154 alressly defined Image 11154 alressly defined Gree Lisket 11154 alressly defined Image 11154 alressly defined um vigues Lisket 11154 alressly defined Image 11154 alressly defined um vigues None Von Lisket um vigues Lisket 11154 alressly defined Image 11154 alressly defined um vigues from matter 1,545 (marge1152+11) um vigues from matter 4,0431 (marge1152+11) um vigues historic matched 40,431 (marge1152+11) um vigues historic matched 40,431 (marge1152+11) um vigues historic None Umatter None None None <tr< td=""><td>International and the second s</td><td>(label ID17 already defined)</td><td></td><td>And</td><td>nov targest</td><td>o meno ar</td></tr<>	International and the second s	(label ID17 already defined)		And	nov targest	o meno ar
maximul columb / 2000	SOR STATED DISTID PSOID.	(label ID16 already defined)		Propert	ies .	ą
Mark matter Gree Mark Market Mark Mark Mark Mark Mark Mark Mark Mark	ISID STATED USITD YSUD	(label ID15 already defined)		🔒 · •	•	
dw Lifet. 107 Artery Writed) None Juends wi "Start disawin". wi Start disawin". Brait 4 of das. wi Start disawin". for aster 5,645 For aster wir Nath Displaces for aster 5,645 For aster More wir Nath Displaces for aster 6,545 For aster More wir Nath Displaces for aster 6,545 For aster More ore college in atter 6,545 For aster More ore college in atter 6,545 For aster More ore college in atter 6,045 For aster More ore college in atter 6,045 For aster More More ore college in atter 6,045 For aster None Wore core core core in atter 10 More 10 More core core core core in aster	save, replace	(label ID14 already defined)		4 Varia	ales	
um ("Digen dide anyphic.") # mili i di dise inter inte	clear	(label ibis aiready defined)				
suf13h1D05h2. not 1.54 for many, mey 13/A1D05h2. for hasier 6.64 6.64 for many, mey 0.2. for many, mey 0.2. for many, mey 0.2. for many 0.	use "G\panel data analysis\	Result	t of obs			
mege 113 MATD 03DD- merge 113 MATD 03DD- ws 5MATD 03DD 75 Mub. ws 5MATD 03DD 75 Mub. ws 5MATD 03DD 75 Mub. mege 115 MATD 03DD. not nat ched f cm wsig t cm vsig t cm v	sort STATEID PSUID DISTID					
name_map_memb2 workMDD00100 ore or cytomatika_map_lensity.ensit	merge 1:1 STATEID DISTID	not matched	9,045			
eschall DB0170 PUID. eschall DB0170 PUID. eschall DB0170 PUID. from sizing 2.134 (perspecialize=1) eschall eschall DB0170 PUID.	rename (merge (mergeR2)).	from master	6,911 (_mergeR1R2==1)			
uar Ma2/s4 (owr uw VStauSNA Ankle. natcled (0,911 (mrg4SH2+1)) Pom Los (mrg bit shown (mrg2 might bit bit bit) mege Lin SMED DSTD. Non Non Non Command Command Non Non Command Non Non Non Command Non Non Non Command Non Non Non	sort STATEID DISTID PSUID	from using	2,134 (_mergeR1R2==2)	Notes		
desc matched (6,558 (mrcg4132+1)) Innum Innum Mohl da Ion Innum Mohl da Ion	save ihds2_link			4 Data		
ue "Gibuiti% Asila. mege Les SAILD DOTD. Command Command Silver Units	clear	matched	40,018 (_mergeR1R2==3)	= files	me bhah	
mega In STATED DSTD	use 'G\Stata15\bb ibdk1.dt		b	1 miles		
command 0 00000100 0000 Command 0 00000100 0000 Very 12000 50000 00000 Very 12000 Society 12000	menne 1 m STATEID DISTID.			in Notes		
Command (193) Command (193) Ser (1943) Merrory (204 Serte (1943)	inege thi shirds outlo.			• note:		
Command Occurrence end Set 446.04 Merroy 1204 Sortet p						
Start Halk, Mi Menory Soried by		Command		• Cose		
Merory 310M Sonted by				Size	746.8	200
					_	_

I think merge know the first line to be. So, it is the link to be given. So, that is basically here using file that has to be first deleted. Then now this is not there it is here. The link we have created, alright. Basically, after using the path name correctly, we observed that we have successfully matched the correct number that is of 40018. And from the Masters, those who are not included.

That is of 6911. And from the using file, that is basically the merge file, the replacement numbers that is up to 2134 we guided from the beginning. Those are not included. So, the net number to be defined as the merge is the panel information is of 40018.

(Refer Slide Time: 38:01)



And we will clarify further. So, basically when we have three indicators, that one code one, 1, 2 and 3 are there. So, what we are interested, only for 3, because three are the common IDs, common information in both round. So, we can draw for one and two because these are no longer important for the analysis. So far as panel is concerned, we can draw and can proceed accordingly.

(Refer Slide Time: 38:34)



So, you must have noticed in the case of first merging, that is the link with IHDS-2. We have used with one is to one merging dividing while in the second one, we merge the IHDS-1 with the

combined link file. So, we applied one to one multiple file. That is merging technique. In first case, information about all the households were given in linking file. In the second case, it's multiple merging because we know some of the IHDS-1 households are split into two or more. And they share same household ID from IHDS-1.

Since there are more numbers and because of split we applied one to m as the merging technique. so the 2 common the total number of household would be in this case is 6911, 2134, 40018 that is of 49063 households so far as the total number is concerned. But out of that 40018 is the final number and they are interpreting both the rounds. Alright.

(Refer Slide Time: 39:44)



So, we can keep that 3. You can do this experiment on your own. We are not repeating. And next, you can simply save with the panel HHID, and that has been already guided and will show while we are using that data for further analysis. So, coming to the dataset which we have already shown for your analysis to make a panel version. You can also click on this link to get for those datasets directly.

(Refer Slide Time: 40:17)



And readily available provided by the ICPSR and they are providing in wide panel, long panel, pooled cross section and appended form as well.

(Refer Slide Time: 40:31)

Name	Size	Preview	Download
🖻 DS1 Individuals - Appended Cross-sections (Public-Use)	3 GB	۲	٤
B DS2 Individuals - Appended Cross-sections (Restricted-Use)	0 bytes		
🖻 DS3 Individuals - Wide Panel (Public-Use)	2 GB	۲	٤
😂 DS4 Individuals - Wide Panel (Restricted-Use)	0 bytes		
😂 DS5 Individuals - Long Panel (Public-Use)	2 GB	۲	¥
😂 DS6 Individuals - Long Panel (Restricted-Use)	0 bytes		
rightarrow DS7 Households - Appended Cross-sections (Public-Use)	1 GB	۲	¥
DS8 Households - Appended Cross-sections (Restricted-Use)	0 bytes		
🖻 DS9 Households - Wide Panel (Public-Use)	924 MB	۲	*
\bowtie DS10 Households - Wide Panel (Restricted-Use)	0 bytes		
😂 DS11 Households - Long Panel (Public-Use)	865 MB	۲	¥
DS12 Households - Long Panel (Restricted-Use)	0 bytes		

For whichever the requirement you have you can download and start using it. So, these are all information given for clarity. You can use it for better explanation further alright. So, I think we need to proceed for the use of all those data that we have already converted, we have made panel and in the next class we will be suggesting you the exact technique of deriving result based on

the panel. So, that is all for today's lecture. Next lecture will be purely understanding and analysis of the panel data. Thank you so much.