Handling Large-Scale Unit Level Data Using Stata Professor Pratap C. Mohanty Department of Humanities and Social Sciences, Indian Institute of Technology, Roorkee Lecture 40 Analysis of Panel Data in Stata

Welcome, friends, once again to the very last lecture of this module on Analysis of Panel Data in Stata. Since you have been following the lectures very seriously, I can infer that you can able to surely interpret the panel data in this format very correctly. And so far in last two lectures we have defined and constructed the panel data carefully and we mentioned using the IHDS both the rounds of the data. And in this final lecture, we will be analysing the panel data.

(Refer Slide Time: 01:20)

INTRODUCTION
The focus of this lecture will be on analysis of unit level data in stata.
For this purpose we are using IHDS household panel dataset.
IHDS is the only longitudinal survey available in India.
There are two waves of this survey IHDS-I (2004-05) and IHDS- II (2011-12).
□ The total number of household surveyed in both the rounds are 40,018.
💿 swayanı 👲 2

So, the focus here will be on the unit level data in Stata and for this purpose, we are using IHDS household panel data and IHDS is the only longitudinal data, as I already mentioned. There are two waves that is IHDS-I and IHDS-II. The total number of household survey in both the rounds is of 40,018 households.

(Refer Slide Time: 01:38)



We are going to analyse the aspects such as we wish to know the effect of work location, then paid workers, religion or caste on performance of the family business. So, in this data, suppose we wanted to analyse the performance of the family business through their religion, caste, whether they are paid workers or their location. So, we are sticking to the performance of the workers since that is not directly observed, we take a proxy variable called a gross receipt of the household business in its log form.

So, gross receipt in absolute number is there. So we have converted into a logarithmic transformation of it. So, the independent variable of interest for us is, as I already mentioned, business location that is maybe home based or fixed place or moving type or paid workers are like, whether they are being paid for their business and it is assume that higher the amount paid more the workers are hired as informational number of workers hired are not provided directly.

(Refer Slide Time: 02:54)



Relating to caste or religion, the variable is categorical in nature and categorized in six numbers or six categories, forward caste, OBC, Dalits, Adivasis, Muslims, Christians, Sikhs and Jain are combined together. The idea behind taking this variable as independent is understanding the effect of culture on performance of the business.

To be noted, the purpose of this lecture is to show how we use various data analysis and their commands. It does not give the whole research process, such as the verification of the assumption of the models, their diagnostics and various follow up analysis usually we do in our earlier lectures, but here we only give you how to use panel.

(Refer Slide Time: 03:58)

. d grossR	ceipts cast	eReligion 1	businessLoca	tion paidWorker
variable na	storage ame type	display format	value label	variable label
grossReceij casteRelig: businessLo paidWorker	ots double ion int cat~n int double	\$11.2f \$23.0f \$13.0f \$10.2f	GROUPSE NF7	HQ14 0.3 Busns1: Cross receipts HQ3 1.13-15 Caste/religion 6cats HQ14 0.7 Busns1: Work place HQ14 0.4c Busns1: Paid workers
				No variable in string form, regression can be run.

Basically, we need to check here that no variables are in string form. That is very important. And then only regression can be run. So, we need to check through the describe command.

(Refer Slide Time: 04:13)



Some of the basics of the panel data analysis is that to use panel data commands in Stata, we need to declare cross-sectional household id, time series survey id. Here it is that survey id, in the name survey or the variable to be with time series component. We need to specify to Stata which variable is cross-sectional and which one is time series. Most of the analysis for panel data in

Stata can be done with xt command. Xt is highlighted here. Xt command is very relevant in most of the panel analysis.

So, xt commands require data to be a long form to make it very clear that the data has to be a long form. Long and wide, we have already mentioned, but we will also show it with the help of data. This means that each observation is a pair of individual and time. Each observation must be given with the time component with the individual information.

(Refer Slide Time: 05:43)

🗟 _ swayam 🛞



- □ In this format, the repeated measurements are distinct variables across the horizontal of each record.
- □ In the last lecture, we have constructed the panel in wide format to match the households interviewed in both the waves.
- But for panel analysis we have to convert them to long form.
- □ Conversion of wide form to long form and long form to wide form is possible in stata with reshape command.

While analysing repeated measures the layout of the data is often in wide format. In this format the repeated measurements are distinct variables across horizontal of each record. Basically, each repeated information are given in horizontal series or horizontal entries.

So, in the last lecture, we have constructed the panel in wide format to match the household interviewed in both the waves. But for panel analysis we have to convert them to a long form for analysis. The conversion of wide form to long and long to wide form is very much possible in Stata with a command called reshape. Reshape is highlighted in blue colour. So, we will do it right now.

(Refer Slide Time: 06:17)

Reshaping the data:
If you want to convert your data from wide format to long format, reshape long command will be used.
For long to wide conversion, reshape wide command will be used.
Essentials for reshape long command:
An id variable
Repeated measures variable – set of variables on which you have repeated information.
Give name to a new variable that index the occasion of the repeated measures. (time variable)
🖲 swajan 👲 8

So, reshaping the data is necessary if we want to convert your data from wide format to long format. So, we will simply give reshape long command, reshape long and it will be converted very clearly. Similarly, if long is available, we need to convert it to wide, so simply we will give reshape wide as the command. So, essential for reshape long command is that an id variable. Those must be discussed like id variables, repeated measures variable, set of variables on which you have repeated information and give name to a new variable that is index the occasion of the repeated measures that is time component. So, time component and id information has to be very clearly specified.

(Refer Slide Time: 07:08)



, тах , тах										
v Tax	📑 🖪 O · O									
									 Variables 	,
ler commands have		(R)							A Liter variables	
er commands nere	1- 1-	// <u>/</u> ,	Conversion 1	005-2012 StataCore	110				Nume	label.
ommand jc	Statistics/Dat	a Analysis	StataCorp	985-2017 Statacorp	LDC				Name	LIDEI
e 'G/panel data analysis),		• 10101/010	4905 Lakeway	y Drive					Id .	group(STATEID
	Special Edi	tion	College Stat	tion, Texas 77845 U	SA				castenergion1	L casteneigion
			800-STATA-PO	C http://www	.stata.com				grosswecerpist	1 grossieceipis
			979-696-460	0 stata@stat	a.con				paroworkers	1 paiowoner
			111-010-400	1 (10A)					IGVAL N	1 1014
	24-student frata	ish propriet in							casteBelinion?	2 casteReligion
	Derial na	#Dec #115162456	114						atoxReceipts2	2 mmsReceipts
	License	d to: IIT Roorke	ee						paidWorker2	2 paidWorker
									businessLoca.	2 businessLoca
	Notes:								IGVA2	2 IGVA
	1. Unicod	He is supported; s	see help unicode_a	advice.						
	2. Maximu	n number of varia	ables is set to 5	000; see help set_m	axvar.					
	. use "G:\panel	data analysis\exa	ample widepanel.dt	ta"					Properties	
	(India Human Dev	elopment Survey F	Panel (INDS, INDS-	-11), 2005, 2011-20	12, Househol	d)			Properties	
									4 Variables	
	· · · · · · · · · · · · · · · · · · ·								Name	
									Type	
									Format	
									Value label	
									Notes	
									4 Data	
									 Filename 	example_wid
									Label	India Humar
									 Notes 	
	(series)								Variables	
	Command								Observations	
									Size	1.58M
										-4
										10
sit View Data Tools	Þ									
id[1]										
id casteReligi	anl grossRecei-1	paidWorkerl busine	essloc-1 10VA1	casteReligion2	grossRecei-2	paidWorker2	businessLoc+2	10%82	* Variables	
id casteReligi	ani grossRecei-i tuslim 6	paidWorkerl busine	essioc-1 10VA1	casteReligion2 .	grossRecei-2	paidWorker2	businessLoc+2	10%82	Variables	
id casteReligi	ani grossHecei-1 Guslim 6 . Luslim 6 .	paidWorkerl busine	essioc-1 10VA1 	casteReligion2 Muslim 6	grossRecei-2	paidHorker2	businessloc+2	10792	Variables	here
id casteReligi 1 9 2 9 3 9	anl grossRecei-1 Guslim 6 . Guslim 6 . Juslim 6 .	paidWorkeri busine	essioc-1 10VA1 	casteReligion2 Muslim 6 Muslim 6	grossRecei-2	paidMorker2	businessloc+2	10782	Variables Variables Variables Variables Variables Variables Variables	here Label
id castelligi 1 3 2 3 3 4	ioni grossRecei-i fusiim 6 . fusiim 6 . fusiim 6 . lusiim 6 .	. paidWorkeri busine	essloc=1 10VA1	casteHeligion2 Musium 6 Musium 6 Musium 6	grossRecei-2	paidHorker2	businessior-2	10782	Variables Variables Variables Variables View variab	here Label group(STATEID
10 castelligi 1 1 1 2 3 3 3 6 3 5 3	ioni grossRecei-i fuslim 6 - fuslim 6 - fuslim 6 - fuslim 6 - huslim 6 -	paidWorkeri busine	Esploc-1 107A1 fixed 2 .	casteReligion2	grossRece1-2	paidWorker2	businessloc-2	10%2	Variables Variab	here Label group(STATEID 1 casteReligion 1 crossReceipti
14 casteRelig1 1 1 2 3 3 5 4 3 5 5 6 5 6 5	anl grosskeci-1 taslin 6 - taslin 6 - taslin 6 - taslin 6 - taslin 6 - taslin 6 - taslin 6 -	peidWorkeri busine	Esplor-1 107A1 fixed 2 . fixed 2 . fixed 2 10.94139	casteReligion2 Musilm 6 Musilm 6 Musilm 6 Musilm 6	grossRecei-2	paidHorker2	businessior-2	1,0%),1	Variables Name Name Name Sine	here Label group(STATEID 1 casteRetigion 1 grossReceiptr 1 pastWorker
14 casteRelig1 1 1 2 3 3 5 4 5 5 5 6 5 7 5	Ini groadheoli-1 Nailin é . Cailin é . Cailin é . Cailin é . Cailin é . Stétés.52 Nailin é .	i paldNorberi busine . Other	Essloc-1 10VA1 fixed 2 fixed 2	casteBeligion2 Muslim 6 Muslim 6 Muslim 6 Muslim 6 Muslim 6 Muslim 6	grossRecei-2	paidHorker2	businessior-2	10%2	Variables Name Name S Name S of C of	here Label I casteReligion I grossReceipti I paidWorker I buinesslocat
id castellizit 1 1 2 2 3 3 4 3 5 2 6 3 7 3 8 3	anl grossHecei-1 Nuslin é - Cuslin é - Cuslin é - Cuslin é - Cuslin é - Cuslin é - Cuslin é - Luslin é -	i peldNorWer1 busine	Emploc-1 10%11	carteleligios2 Nusian 6 Nusian 6 Nusian 6 Nusian 6 Nusian 6 Nusian 6 Nusian 6	gzosskecei-2	paidHorker2	Businesslor-2	10%3	Variables If the variables Name is a casekeipon orsikecept1 pastWorker1 whatesticac. WAIA	here Label I casteReligion I grossReceptt I gorsReceptt I businessi.ccat I SVA
id castellist 1 1 2 1 3 2 4 3 5 2 6 3 7 2 8 3 9 3	ani quassincei-i Nuslin é Guslin é Guslin é Guslin é Guslin é Guslin é Guslin é Juslin é Juslin é	i paidWorberi busine	essice-1 10%11 	casteRtigios2 Navias Navias Navias Navias Navias Navias Navias Navias Navias	gtossheeti-2	paidKorker2	Businestor-2	10%2	Variables Variables P. Name P. Si of C. casekenpont P. poskeceptit P. po	here Label I casteReligion I grossReceipti I pad/Worker I businessi.ccal I SVA 2 casteReligion
14 castelligit 1 1 2 1 3 3 4 3 5 3 6 3 7 3 8 3 9 3 9 3 10 3	- ani greadWeel-1 Musine 4 - darline 4 - harline 6 -	peldWorker2 bugine . Other	Essloc-1 10%1 	casteletigics2 Mulia 4 Mulia 4 Mulia 4 Mulia 4 Mulia 4 Mulia 4 Mulia 4	grossRecti-2	paidHorker2	Businestor-2	ເຫນ	Variables Variab	here Label group(STATEID I castelReligion I grossReceipti I paidWorker I businesLocal I SOVA 2 castelReligion 2 grossReceipti
16 castelligit 2 1 3 3 4 3 5 3 6 3 7 3 8 3 9 3 10 3 11 3		i paidWorker: busine	Essloc-1 10VA1 faxed 2 faxed 2 Some 1 10.4354 Faxed 2	casteleigins Nuile Nuile Nuile Nuile Nuile Nuile Nuile Nuile	grosskeet-2	paidHorker2	Businestor-2	1992	Variables Name Name Varia	here Label group(STATEID 1 casteReligion 1 grossReceipt 1 businessLocal 1 IGVA 2 casteReligion 2 grossReceipt 2 grossReceipt 2 grossReceipt
16 caseship 1 1 2 1 3 3 4 3 5 3 6 3 7 3 8 3 9 3 10 3 11 3 12 3		 pil@vz@wri bupidwi cher cher 0.00 cher 0.00 0.00 cher 	Essloc-1 12VA1 	Castelleigios2 Masia 6 Masia 6	grossRecei-2	paid%utker2	Businesior-2	19%)	Vanables (Increasedine A Increasedine Casteleripoil Castelerip	here Label group(STATEID L asteReteigon L pardWorker 1 businessLocat 1 KVA 2 casteReteigon 2 grossReceipti 2 pardWorker 2 businessLocat
14 certologia 1 1 2 1 3 1 4 2 5 2 6 2 7 2 8 2 9 2 10 2 11 2 12 2 13 2	Image: Second	i paldHotheti buallo 	eseloc-1 10%11 	Castebeligios2 Notine 4 Notine 4 Notine 4 Notine 4 Notine 4 Notine 4 Notine 4 Notine 4 Notine 4 Notine 4	grosskecti-2	paidWorker2	Businesior-2	1094) - - - - - - - - - - - - - - - - - - -	Vanables Vanables Vanables Vanables Vane Vanables Vane Vanables Vane Vanables Vanebles Vanebl	here Label group(STATEID 1 casteReigon 2 possReceipt 2 pastReceipt 2 pas
iii carstelling 2 1 3 1 4 3 5 3 6 3 7 3 8 3 9 3 10 3 11 3 12 3 13 3 14 3 15 3		i paidWatheri bugalo 	exion-1 10%11 fixed 2 10.94199 Mumo 1 10.49214 Mumo 1 10.49214 fixed 2 	Casteleigion2 Marine 6 Marine 6	grosskeet2	paidWorWer2	Businessior=2	1594	Variables Variables	here Label group(STATEID 1 casteReligion 1 grossReceipti 1 pas/Worker 1 businessLocat 2 casteReligion 2 grossReceipti 2 businessLocat 2 businessLocat 2 SWA
iii carstelling 2 1 3 1 3 2 3 2 4 2 5 2 6 2 7 2 8 2 9 2 10 2 11 2 12 3 14 2 15 2 16 2			estico-1 10%11	Castekkişteni Rusine & Rusine & Rusine & Rusine & Rusine & Rusine & Rusine & Rusine	grossfeets-2	paidHorber2	Businestor-2	1002	Variables Variables Variables Varia Varia Variables Variable	here Label group(STATEID L casteReligion L grossRecepti D usinessLocal D usinessLocal 2 casteReligion 2 grossReceipti 2 padWorker 2 businessLocal 2 sGVA
iii carstelling 1 1 2 1 3 1 4 3 5 3 6 3 7 3 8 3 10 3 11 3 12 3 13 3 14 3 15 3 16 3 17 3 18 3 19 3 10 3		pai/Herberi bushe 	esion-1 12701 faxes 2 faxes 2 faxes 2 	carteleigent Noria e Noria e	grossRecti-2 - - - - - - - - - - - - - - - - - -	paidHorber2	Businestor-2	19982 	Variables (for examples (for examp	here Label proup(STATEID L cateReligion 1 grosiReceipti 1 businessLocal 2 GateReligion 2 grossReceipti 2 possReceipti 2 pastWorker 2 businessLocal 2 IGVA
64 carseletion 2 1 3 1 4 2 5 2 6 2 7 2 8 2 9 2 10 2 11 2 12 2 14 2 15 2 14 2 15 2 16 2 17 2		galdheteri busho , , , , , , , , , , , , ,	199520-1 10995 Taref 2	custeltigics Notice Not	grasskect-2	paidHor2Her2	BusinesSor-2	1011	Variables Variab	here Label group(STATEID 1 parkleigot 1 parkleigot 1 parkleigot 2 tastekeigot 2 businessLocat 2 parkleigion 2 park
64 caracteristic 2 1 3 1 4 2 5 1 6 2 7 2 8 2 9 2 10 2 11 2 12 2 13 2 14 2 15 2 16 2 17 2 18 2 19 2 10 2 11 2 12 3 13 3 14 3 15 3 16 3 17 3 18 3 19 3 10 3	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	pullPrenet: builde	ndard 1094	conditional Realist & Realist & Real	grosskect-2	pa10%stbsz2	Businession-2	19982 	Vanables Van	here Label proug(STATEID 1 cateReligion 1 partWorker 1 businesilocat 1 IGVA 2 grossReceipti 2 partWorker 2 businesilocat 2 IGVA 2 ISVA
id carstelling 1 1 2 1 3 1 4 2 7 2 8 2 9 3 10 2 11 2 12 2 14 2 15 2 16 2 17 2 18 2 14 2 20 2 20 2	Image Construction Notation 4 -	puldhemet kuase 	natarei 1996) 	Line (1997)	grossRect2 - - - - - - - - - - - - - - - - - -	paidHother2	Businestor-2	1093	Variables Variab	here Label group(STATED) LateRelegion L grossReceipti D businessLocal L KVA 2 casteRelegion 2 postReceipti 2 padVloker 2 padVloker 2 businessLocal 2 NOVA
iii outstain 2 1 3 1 4 1 5 2 7 2 8 2 9 2 10 2 11 2 12 2 13 2 14 2 15 2 16 2 17 2 18 2 19 2 10 2 11 2 12 3 13 2 14 3 15 2 16 3 17 3 18 3 19 3 10 3 11 3 12 3 13 3 14 3 15 3 16 3	Image Image <thimage< th=""> <thi< td=""><td>paid/ventor busission <td>1998-1 1998 </td><td>Longitabres 4. Airco 4. Airco 4.</td><td>groasheet2</td><td>paidHorker2</td><td>Businession-2</td><td>19932 </td><td>Vanables Vanabl</td><td>here Label Josef Label Label LasteRetigion grossRecepti JosinesLocal Josef LastRetigion grossRecepti 2 padWorker 2 businesLocal 2 NOVA</td></td></thi<></thimage<>	paid/ventor busission <td>1998-1 1998 </td> <td>Longitabres 4. Airco 4. Airco 4.</td> <td>groasheet2</td> <td>paidHorker2</td> <td>Businession-2</td> <td>19932 </td> <td>Vanables Vanabl</td> <td>here Label Josef Label Label LasteRetigion grossRecepti JosinesLocal Josef LastRetigion grossRecepti 2 padWorker 2 businesLocal 2 NOVA</td>	1998-1 1998 	Longitabres 4. Airco 4.	groasheet2	paidHorker2	Businession-2	19932 	Vanables Vanabl	here Label Josef Label Label LasteRetigion grossRecepti JosinesLocal Josef LastRetigion grossRecepti 2 padWorker 2 businesLocal 2 NOVA
Her Cartolization 2 1 3 1 4 1 5 1 6 1 7 1 8 1 9 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1 1		puldhemet baalse 	Eastard 1998) 	Sequences and a second	grossRect2 - - - - - - - - - - - - - - - - - -	PAL96178427	Businesico-2	10%)	Variables Variables Variables Variance Variance Variance Variables Var	tere Label Label Lastekeispon Japosläkeepis Japosläkeepis Japosläkeepis Danessiscata Danessiscat
Her Caracteristica 2 1 2 1 3 1 4 1 5 1 6 1 7 1 8 1 9 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1		pil@verteti budase 	UNITES 1994	Sequelabores 	200000.00 	943.0%121612 	Businession-2 - - - - - - - - - - - - - - - - - -	19792	Vanables Vanabl	here Labet groupSTATED I gootkeepts I gootkeepts J pointeepts J pointe
44 Caracteristic 2 1 2 1 3 1 4 1 5 1 6 1 7 1 8 1 9 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1	Image Image transa	PLIBURES IN INTER 	1996-1 1996. 1996 - 1996. 1996 - 2 1996 - 2	Languistana 4 atuan 5 atuan 5 atuan 5 atuan 5 atuan 6 atuan 6 atuan 8 atuan	grossRect2 - - - - - - - - - - - - - - - - - -	943.0%019427 	Businestion-2	10%	Variables Variables Variables Variant Variant Variables Vari	Ince Label Jackel Jogosheepin Jackel Jogosheepin Jackel Ja
H Constraint 1 1 2 1 3 1 4 1 5 1 6 1 7 1 8 1 9 1 10 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1	Image Image <th< td=""><td>pil@vriteti budase </td><td>uninesi 1998. </td><td>Longitships </td><td>9000.000 100000.0000 10000000000</td><td>983.0%42%422</td><td>Businessor-2 - - - - - - - - - - - - - - - - - -</td><td>19792</td><td>Vanables Vanables Van</td><td>here Labei groupSLATED La catelifique providences La partiNoter La particular 2 cateletigue 2 catele</td></th<>	pil@vriteti budase 	uninesi 1998. 	Longitships 	9000.000 100000.0000 10000000000	983.0%42%422	Businessor-2 - - - - - - - - - - - - - - - - - -	19792	Vanables Van	here Labei groupSLATED La catelifique providences La partiNoter La particular 2 cateletigue 2 catele
44 Cartolizing 2 1 2 1 3 1 4 1 5 1 6 1 7 1 8 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 12 1		PaliPartiti badas 	nonier-1 1994) filmet 2 1. filmet 2 1. fi	Languistana 4 sina 4 sina 5 sina 5 sina 5 sina 4 sina 1 sina 1 sina	grassheeti-2 - - - - - - - - - - - - - - - - - -	pa.dHuter.2	kutimitari Otor fast J Otor fast J Otor fast J	18932 	Vanables Vanabl	here Label genesiSTATEED 1 castelligion 1 castelligion 2 partitioner 1 kiva 2 partitioner 2 partitio
H Constraint 1 1 2 1 3 1 4 1 5 1 6 1 7 1 8 1 9 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1	J J Dista 0 Dista 4 Dista 1 Dista	padmenti badan 	Tomi 2 1994: Tomi 2 1.0101 Ram 1 1.0104 Ram 1 1.0104 Tomi 2 1.0104 Ram 1 1.0104 Tomi 2 1.0	Sequelations 4 along 4 along	2-1998/44459 	pa.00x2x22	businesides d	19792	Vanables Vanabl	tee Läket geosefstätet japotkeest
44 Cartolizing 2 1 3 1 4 1 5 1 6 1 7 1 8 1 9 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 12 1		PaliPartetti badasi 	19941-1 19941 1994 1 1994 2 1994 2	Languistenes 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	graadee;	9430012802 	Rusinsteed	18792	Vanables Vanabl	tere Laket Depositency Laceteringon Depositency Journal Depositency Depositenc
H Constraint 1 1 2 1 3 1 4 1 5 1 6 1 7 1 8 1 9 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1	J J J 000000000000000000000000000000000000	padmenti badad 	Fore 2	Sequences		9410912122 - - - - - - - - - - - - - - - - -	Restoration 2	19792 	Vanables Vanables Vanables Vane	terr Laket geosoffat(ED geosoffat(ED geosoffate) Lanethrigon Janthoner Lanethrigon Janthoner Jan
44 Caracteristica 2 1 2 1 3 1 4 1 5 1 6 1 7 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 20 1 21 1 22 1 23 1 24 1 25 1 26 1 27 1 28 1 29 1 20 2 21 1 22 1 23 1 24 1 25 1 26 1		PatiBratests backst 	1994(-) 1994) 1994 19	Languishese 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4		94100-12421 	Residenties	1992	Vanables Vanabl	Here Label II cateletigion geotiknegi Jourisetta 2 cateletigion 2 postbiore 2
H Constraint 2 1 2 1 4 1 5 1 6 1 7 1 8 1 9 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1 <tr tr=""></tr>	Image Image <th< td=""><td>padmenti badad </td><td>randor-1 1994: Tanet 2 13-0139 Road 2 13-0139 Road 1 13-0139 Tanet 2 13-0139 Road 1 13-0139 Tanet 2 13-0139 Tanet 2 13-0139 Road 1 13-0139 Road 1 13-0139</td><td>Langithten 4 along 4 along</td><td></td><td>pa160/12021 </td><td>Buildendard</td><td>1893</td><td>Variables Variables</td><td>here Label LaceNetipoys LaceNet</td></th<>	padmenti badad 	randor-1 1994: Tanet 2 13-0139 Road 2 13-0139 Road 1 13-0139 Tanet 2 13-0139 Road 1 13-0139 Tanet 2 13-0139 Tanet 2 13-0139 Road 1 13-0139 Road 1 13-0139	Langithten 4 along 4 along		pa160/12021 	Buildendard	1893	Variables Variables	here Label LaceNetipoys LaceNet
bit Catalogue 2 1 2 1 3 1 4 1 5 1 6 1 7 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 13 1 14 1 15 1 16 1	Image Image angina 4 angina 4 <	PLIBURENCE INALMA PLIBURENCE INALMA COMPANY C	none	Lenguishese 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	201404043 	Pat/80/196/1 - - - - - - - - - - - - -	Bolititude-2	19792	Vanables Vanabl	her Liber I catherligen geosthereigt J catherligen J cathe

Like in the example we will also provide this information to you. We have an example wide panel data that we converted in the last class also. But we are discussing once again with the name example wide panel data. We will simply use that data right now, example wide panel data, and from there we will discuss like panel data example wide panel.

So, this has been opened. You can see the data is given. I will also interpret here. This is the variable and label is very clearly given at the right hand side. Variable name like caste, religion with one, gross receipt one and period one, then it is caste, religion and period two, gross receipt period two, paid worker in period two. So, those are given very clearly and you can check the

data also. And since this is a wide panel, because the repeated information related to caste, religion, it is given in horizontal order. But Stata requires for the analysis, we generally require a long format data we have to convert it.

So, this is what the data created. The variables are named with a suffix where 1 represents time period 1 and 2 represents time period 2. And like some variables are time invariant, like caste of the person generally of time invariant, but largely most of the variables are changing over time. So, better to rename with different variable because they are changing nature of those variables.

(Refer Slide Time: 09:21)

 Correshape this format in long: cashape long casteReligion grossReceipts paidWorker paidwork			
<complex-block><pre>eshape long casteReligion grossReceipts paidWorker businessLocation IGVA, i(id) j(SURVEY)</pre> The variable names nothing but the prefixes from the previous repeated variables. DURVEY is the occasion variable created. Vive variable names nothing but the prefixes from the previous repeated variables. DURVEY is the occasion variable created. Vive variable names nothing but the prefixes from the previous repeated variables. DURVEY is the occasion variable created. Vive variable names nothing but the prefixes from the previous repeated variables. DURVEY is the occasion variable created. Vive variable names nothing but the prefixes from the previous variable names nothing but the prefixes from the previous variable created but the previous variable castered but the previous variable names nothing but the previous from the previous fro</complex-block>	To re	eshape this format in long:	
<pre>seshape long casteReligion grossReceipts paidWorker businessLocation IGVA, i(id) j(SURVEY)</pre> The variable names nothing but the prefixes from the previous repeated variables. SURVEY is the occasion variable created. SURVEY is the occasion variable created. Www.execution Variables be developed and the prefixes from the number of variables have decreased but the number of observation have increased. Www.execution HQ1417 Businst Work place Www.execution HQ1417		.	
businessLocation IGVA, i(id) j(SURVEY) The variable names nothing but the prefixes from the previous repeated variables. SURVEY is the occasion variable created. Image: state of the state	reshap	e long casteReligion grossReceipts	paidWorker
 The variable names nothing but the prefixes from the previous repeated variables. SURVEY is the occasion variable created. Interview of the variable names nothing but the prefixes from the previous repeated variables. The number of variables the variables created but the number of variables have decreased but the number of observation have increased. 	hucina	sclocation IGVA i/id) i/SLIDVEV	•
<complex-block></complex-block>	DUSITIE	SLOCUTION IGVA, ITU) JSORVET	
 The variable hannes heatings but the previous repeated variables. SURVEY is the occasion variable created. Ivitie the oc	The	variable names nothing but the prefixes	from the
<complex-block></complex-block>	e	invested weighted	
SURVEY is the occasion variable created. Image: Subscription of the operation	prev	lous repeated variables.	
Image: space spac	🗅 SUR	VEY is the occasion variable created.	
id group(STATED DISTID FSUD HHD HHSP. SURVEY HOSS (2003) (HDS2 (2012) catsReligion (HJ13.13) Statter(religion Reas have decreased but the number of observation have increased. id/Work HQ18.8.8 burnt: Bod workers have increased. id/Work HQ18.8.7 burnt: Work place have increased. id/Work HQ18.8.7 burnt: Work place id/Work id/Work HQ19.8.7 burnt: HQ19.9.0.0 burnt: HQ19.9.	- E	Variable Label	
SURVEY - HHOSE (2002) or HHOSE (2012) casteReligion HQ3 1.13-15 Caster/religion 6cass padWorker HQ14 8.4 Burns1: Work place NVA Note: Increased. Note: Incr		id group(STATEID DISTID PSUID HHID HHSP. The number	er of variables
Imposed provide and particular provide and partic		survey IHDS1 (2005) or IHDS2 (2012) casteReligion HO31.13-15 Caste/religion ficats	eased but the
pidWorker HQ148.46 Bunns1: Work place total total </td <td></td> <td>grossReceipts HQ14.8.3 Busns1: Gross receipts number of</td> <td>f observation</td>		grossReceipts HQ14.8.3 Busns1: Gross receipts number of	f observation
bulinesd.ocation HQ14.8.7 Burns1: Work place DVA Note: Image: Im		paidWorker HQ14 8.4c Busns1: Paid workers	observation
Image: control in any properties (see high unit control		businessLocation HQ14 8.7 Busins1: Work place NaVe II	hcreased.
Control Density for White Here Special Edition Special Edition <tr< td=""><td></td><td>IGVA</td><td></td></tr<>		IGVA	
A Grand Market Ver Nove Net	ewayam	A	10
A Guyne Some is Weeke He A guyne Some is Balance String A guyne Some is Balance Some is	Swayain	979 	
Number of dots, 40018 → 80314 → 100018 → 100018	▼ ♣ x stands here sd _K anel data analysish, ong casteReligion _	405 Latency Strive Special Ratten College Station, Franz 7745 USA 80-1787A-N Notesti	Variables T Name Label id groupSTATED SURVY POST (2005) of canteligon HQ1111 STant grotRecets HQ114 STant burdholmer HQ14 & Romit burdholmer HQ14 & Romit
j tariable (i values) → sworr s) tariable (i values) → catathaligen gransfieligi (astabaligen) → catathaligen gransfieligi (astabaligen) → catathaligen possibility (astabaligen) → catathaligen butisestoatisti (astabaligen) → balisestoatisn 1001 1002 → 1008		1. Onicode 19 supported tet and particle approach. 2. Nations matter of wrindled is set to 5000; best help set_marwar. .use "C'apacel data analyzin'example widepeak dis" [Edial Ream Development Survey Mark (Edit, Edit, 2005, 2011-2012, Rosenhold) .example long castabaligion groundwoodspic paidWorker businessiccation 10%, 1(d) (ESWART) (more; j = 1) Data vide → long	itiva Properties
castakilipisti cartakilipisti -> castakilipisti prostakistijisti paikaikapti -> pistakietyti -> busissilicettisi busissilicettisi -> 10001 10002 -> 1000		1. Onicode js supporters tet and particle pa	1014 Properties W voldele Norm
builterslocation > puldwither >> puldwither >> puldwither builterslocation >> builterslocation >> >> 1000.11002 >> 100 >> >> Command 100 . . .		1. Onicode js supporters tet and particle pa	Dix Properties Units Turn Turn Turn
Lonner Lond -> Lond 		1. Onlocal ji Supprise ise anj anicas partos. 2. Nuclean mater of variales is set 5500; bes help set_maras. (us "O'yamai data aniziria targing undersani dat" (India Rama Development Survey Rami (INDS, INDS-17), 2005, 2011-2012, Rosenhold) restaga leng cambalaging groundecipts paidforke businestication 1000, 1(1d) 1(100VET) (Inter: j + 1 2) hts vide → long Ramber of variales 11 → 7 j variable (2 values) → SUNVET (1) variable; 1) variable; 1) variable; 1) cartebulgioni castebulgioni quantarcupti castebulgioni castebulgioni castebulgioni quantarcupti castebulgioni	EXX Properties Washing Variable Tase Variable Variable
. .		1. Goldon 1: apporten iete anj malica partie. 2. Nutions maken of variable is set 5500; bes help set_marks. . use "O jusai dat maipinased available is to 500; bes help set_marks. . use "O jusai dat maipinased is the set of the	Dates
Command at 85% Command at 85% Series 1239 Series 1239 Series 1239 Series 1239 Series 123		1. Goldon 10 septeties for any mainten parties 2. Notions mainten of variable is as to 5000 sets bails set_marker. . es "Ojusai data saiyirkangis undessai da" (folia Maan Beelgeent Survey Reed (IRG, ISG-11), 2005, 2001-2002, Nousehold) . estage leng establighen greatbeligte paidforte boinsesfection 1070, 1(dd)(ISGNUT) Mater of dos. 4001 → long Number of dos. 4001 → boolt Number of dos. 4001 → 50016 Number of variable 1 → 7 yariable (Valen) → 2000T (j) variable (Valen) → source scatebolighen paidforter jatebonet → paidforter paidforter jatebonet → paidforter buinsesfection latorescent → paidforter buinsesfection latorescent → boinsesfection 1070, 10702 → 1070.	DOX Properties
5.47 1230 Weinery GAU Soniel by d'SUNY		 1. Goldon 10 supporters for any mainten parton. 2. Nuclean mained of variable is set to 5000 sets bails set_market. - use "O(passi data aniprirkanegis undepast dat" (folia Rama Beelogenet Survey Fanel (IRG, IRG-11), 2005, 2001-2012, Novehold) - reshes long catcheligies greatebaciegies paidforter buinessicceties 1070, 1(14) ((DWWRT) (http://silen	OXA Properties Properties
Somed by d SURVEY		 i. Okcome in apportence inte sain pairing pairing. Nuclean model of variable is sate 15 5000 pice halp set_mares. use "0 juest data satipitizeneopis yodopair dat" (folds Aman Beelogenet Hurry Fanel (1000, 1000-110, 2000, 2001-2001, Household) reshge log catabaligies groutbackeigts paddente buissatication 1070, 1(14) (f00000) (note; j = 1) Nata vide → long Matter of obs. 40010 → 80016 Matter of obs. 40010 → 80016 Matter of variables 11 → 7 yariable (10 yulies) → 3000000 (catebaligies) catabaligies → catabaligies paddenter jationski → paddenter buissatisetient industriation → buissatisetient buissatisetient industriation → paddenter buissatisetient industriation → buissatisetient 1070, 10702 → 1070. 	BUA Paperties
		 i. Okcode ji soporteti ret saji palitagi paritagi pa	EVA Poporties A A Poporties A A Second Contemport Conte
	N 4	 0.0000 19 Reporters for and particle gamma. Nuclear model of variable is the 10 MHz best high set_marks. Nuclear model of variable is the 10 MHz best high set_marks. Nuclear model of variable is the 10 MHz best high set_marks. Instage long establights providencing particular the businestication 100, 1010 (100000) (note () + 1 2) Atta vide > long model of variable is the 10 > 0 of the 10 min model of variable is the 10 > 0 of the 10 min model of variable is the 10 > 0 of the 10 min model of variable is the 10 > 0 of the 10 min model of variable is the 10 > 0 of the 10 min model of variable is the 10 > 0 of the 10 min model of variable is the 10 > 0 of the 10 min model of variable is the 10 > 0 of the 10 min model of variable is the 10 min mo	Poperties Poperties Poperties Poperties Pore Poperties Pore Poperties Poperti

	Lares												
-	10(4)									 			
	10	JUNYET	casteReligion	grossRecel-s	paldworker	businessLoc-n	1074				- â	Variables	
1		10001 1	NUMBER OF									+ Elter variable	
-1-		10022 2	Nuclear 6									🗹 Name	Label
-i-		10041	Number 6									🗑 id	group(STATEID
-1-	- 1	10001	Number 4									SURVEY	IHDS1 (2005) or
		10041 1	Number 6									CasteReligion	HQ3 1.13-15 Ca
	- í	10001	Number 6									grossReceipts	HQ14 8.3 Busns
1		10001 1	Number 4									PaidWorker	HQ14 8.4c Bush
	- 1	18041 1	Nuslin 6									DusinessLoca	HQ14 8.7 Busns
		18001 1	NUSLIN 6			other tikes r						😨 IGVA	
10		10011	NUMBER 6										
		10001 1	Nuslin 6	56165.54	4.04	VUMP LINES 7	10.34133						
		10041	Number 6										
		10001 1	Number 6										
	ĺ.		Nutrin 4										
		18051 1	NUSLIE 6										
10		18002 2	NUSILE 6									Variables Shaps	hots
		18001 1	NUSLIE 6	338/9.31	0.00	NORE 1	10.43036					Properties	
	2	1002 2	Husing 6									4 Variables	
17	10	10001 1	NUMER &	228/9/21	0.00	FORM 1	10.43036					- variauros	
20	10	10004 4	NUSLIN 4									Label	
21	11	18031 1	NUSILE 6	0.00	0.00	other tixed 2						lune	
	11	18037 7	NUSILE 4										
	14	18051 1	NUPILE 4										
		18032 2	NUSILE 6										
	13	18001 1	Nuplim 6									4 Data	
	13	18002 2	Nuslim 6									Filename	
21	14	18001 1	Nuplim 6									Label	India Human
	14	invel i	Nuslim 6									Notes	
29	15	28091 1	Muslim 6									Variables	
30	15	18092 2	Nuslim 6									Observations	80,036
31	16	18091 1	Huplin 6	*1408.05	0.00	other fixed 2	10.63125					Size	2.21M
					10	Anna Acres 5.					,		

We will reshape it to long format. Since I have already shown that it is available in wide format. So, we will simply reshape it. But reshape with this particular name we have to specify what is our id variable because then only Stata can able to reshape it with id against time. So, we have to define what is our j that is time and id variables, who is the id and who is time. So, we will specify it. So, here will simply take it and will find out with this.

Once we specify it we enter it. We can see that these are information. I want to show it. Look at number of columns have been reduced. But the number of observation has been increased to 80,036. I will show you here. Look at this. Here the id is the observation who has responded 40,008 in total in one round. But since we have put it in long format, in vertical format, this is survey round one. This is for same id, the same person who is responded in survey time period one and time period two. The person three-time period one, and similarly, other information has been saved accordingly.

So, I wanted to just inform here that the number of variables were 11 earlier since it was of wide panel. Since we have made it clubbed into a vertical format, a long format, so the number of variables have been reduced, but the number of observation has been increased to 80,036 from 40,018.

So, we will go back to our PPT. So, the variable names nothing but the prefixes from the previous repeated variables. Survey is the occasion variable created that is the time variable

created. The number of variables have decreased, but the number of observation have increased as I already mentioned. So, here also against to our variable and their labels you can easily mark the difference as compared to the previous wide format data. We will come to the next.



(Refer Slide Time: 12:04)

This is what we have explained.

(Refer Slide Time: 12:08)

PREPARING STATA FOR PANEL DATA
We need to set Stata to handle panel data by using the command xtset or tsset.
Xtset and tsset is used to define cross section (i) and time series (t) variables.
Both commands do the same thing.
These two commands allow use of panel commands and some time series operators.
And they automatically sort the data accordingly.
In panel data cross section variable and time series variable should jointly uniquely identifies the data.
Si Swapan 👰 12

We are preparing Stata for panel data analysis. We need to set Stata to handle panel data by using the command xtset or tsset. These two commands are interchangeable used xtset or tsset. So, xtset or tsset they both gives the same result. These commands are used to define cross section that is i and the time component t variables. We will show it right now. Both commands do the same thing. These two commands allow use of panel commands and sometime series operators and they automatically sort the data accordingly. In panel cross data section, panel data cross section variable and time series variable should jointly uniquely identify the data. We will tell you. (Refer Slide Time: 13:10)



Like in IHDS household long panel data we have multiple identifiers that is state ID, district ID, PSUID, HHID and split ID. We cannot specify multiple identifiers with the xtset command. Xtset command we cannot have multiple identifiers. We have to make it a group. We can generate one id variable from these multiple variables with egen command. Egen command also we discussed earlier, but once again we are discussing that egen command with group function once we mention group with all those five variables it will combine into a particular variable.

So, group maps like egen command is an extension of generate command group with the specific variables, maps the distinct groups of a list, of a variable is to a categorical variable that takes an integer values from 1 to the total number of groups.

(Refer Slide Time: 14:17)



<pre> transmission transmission</pre>	19. T. R. R. 10 11 (0.0		
A manual de duisies is du subjeties is du subjeties. A manual de duisies is du subjeties is dues du subjeties. A manual de duisies is dues du subjeties is dues du subjeties. A manual de duisies is dues du subjeties is dues du subjeties. A manual de duisies is dues du subjeties is dues du subjeties. A manual de duisies du subjeties is dues du subjeties is dues du subjeties. A manual de duisies du subjeties is dues du subjeties. A manual de duisies du subjeties is dues du subjeties is dues du subjeties. A manual de duisies du subjeties is dues du subjeties is dues du subjeties. A manual de duisies du subjeties is dues du subjeties is dues du subjeties is dues du subjeties is dues du subjeties is due subjeties is dues du subjeties. A manual de duisies du subjeties is due subj	▼ # × 1	 Unicode is supported; see help unicode_advice. 	 Variables 	τ.
<pre>vist 0 function function</pre>	mands here	 Maximum number of variables is set to 5000; see help set_maxvar. 	+ Fiter variab	
<pre>market. m</pre>	d _x .use *	G:\panel data analysis\example_widepanel.dta"	Name	Label
<pre>media in the second product of the seco</pre>	anel data analysis)	Human Development Survey Panel (IHDS, IHDS-II), 2005, 2011-2012, Household)	OISTID	HQ0 District code
<pre>are specified a specified</pre>	ong casteReligion .	ane long castaBaligion grossBaceints paidWorker husinessLocation 1008. 1(14) 1(108VEY)	PSUID	HQ0 PSU: village/.
<pre>startic. butch butc</pre>	(note:	j = 1 2)	HHID	Household ID IHD
<pre>bit bit bit bit bit bit bit bit bit bit</pre>	el data analysis\		HHSPUTID	HQ0 Split househ.
<pre>vertice d de . dest : e total presente d de . dest : e total generation de destanting de . dest : e total generation de destanting de . dest : e total de de destanting de destanting de de destanting de de destanting de de destanting de de de de de de de destanting de de de de de de de de de de de de de d</pre>	USTID PSUID - DATA	0108 -> 100g	W12012	2012 weights
bereit of statistics bereit of statistics statistics in the statistics	y Number	of obs. 40018 -> 80036	A SLIRVEY	HOS1 (2005) or L
<pre>private in the interview of the int</pre>	Nunber	of variables 11 -> 7	IDPSU	PSU id unique 6 di
<pre>class set set set set set set set set set s</pre>) varia	ible (2 values) -> SUNVEY	IDHH	Household id, uni.
<pre>sedential statuti : : : : : : : : : : : : : : : : : : :</pre>	,	casteReligion1 casteReligion2 -> casteReligion	WT2005	2012 weights
<pre>building building buildin</pre>		grossReceipts1 grossReceipts2 -> grossReceipts	FWT2005	Integer 2012 weig
<pre>international international internation</pre>	b	paidWorkerl paidWorker2 -> paidWorker	REGION	HH0 states group
<pre> ciar int of the inter subjecting part of of (the inter subjecting part of of (the inter subjecting part of of (the inter subjecting part of of</pre>	545	10VA1 10VA2 -> 10VA	STATEID2	State codes, colla.
<pre>ctar int final status brighter int shift into the shift into</pre>			Properties	
<pre>class (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011-2012, Romanojan (total same bevelopest barry has 1005, 1001-101, 2005, 2011</pre>		6	8 * *	
<pre>ins - C (post data subjirtus (post data) ins data (post data subjirtus (post data) ins data (post data subjirtus (post data) ins data (post data subjirtus (post data subjirtus (post data) ins data (post data) ins d</pre>	. clear		Variables	
the serie descriptions berry freed (1985, 1985-197), 2007, 2017-2027, Bouehold) - isis DATED DISTO FROD MED MEDITID DAVIT - serie is + group (INTELD DISTO FROD MED MEDITID) - isis is 4 SOUTH - serie is + group (INTELD DISTO FROD MED MEDITID) - isis is 4 SOUTH - serie is + group (INTELD DISTO FROD MED MEDITID) - isis is 4 SOUTH - serie is + group (INTELD DISTO FROD MEDITID) - isis is 4 SOUTH - serie is + group (INTELD DISTO FROD MEDITID) - isis is 4 SOUTH - serie is + group (INTELD DISTO FROD MEDI MEDITID) - serie is + group (INTELD DISTO FROD MEDITID DISTO FROD MEDITID) - serie is + group (INTELD DISTO FROD MEDITID DISTO FROD MEDITID) - serie is + group (INTELD DISTO FROD MEDITID DISTO FROD MEDITID) - serie is + group (INTELD DISTO FROD MEDITID DISTO FROD MEDITID) - serie is + group (INTELD DISTO FROD MEDITID DISTO FROD MEDITID) - serie is + group (INTELD DISTO FROD MEDITID DISTO FROD MEDITID) - serie is + group (INTELD DISTO FROD MEDITID DISTO FROD MEDITID) - serie is + group (INTELD DISTO FROD MEDITID DISTO	. use -	G:\panel data analysis\## long panel.dta"	Name	SURVEY
<pre>init initial properties prop</pre>	(India	Human Development Survey Panel (IHDS, IHDS-II), 2005, 2011-2012, Household)	Label	0HDS1 (2005) o
<pre>ind STATUS DISTOR FULD MAD MERITIES SAVET</pre>			Type	
<pre>+epu id + group (TSATED DITOR FUDD NED NEFLITED)</pre>	. isid	STATEID DISTID FSUID KHID KHIPLITID SURVEY	Format	
<pre>even is + group (SENTED DISTO FUED REFILITO)</pre>			Value label	
<pre>id de source in de de source de in de source de in de source de in de source de in de source de in de source in de so</pre>	. egen	id = group(STATEID DISTID PSUID HHID HHEPLITID)	Notes	
<pre>bill d J JOHT bill d JOHT bill d JOHT bill d J JOHT b</pre>			Elename	
<pre>intermediate intermediate intermediate</pre>	. 1910	1d SUMVET	Label	
A Construction Constructi			 Notes 	
Control Control			Variables	581
A second seco			Observations	80,036
Image: 10	Comman	0	Size	183.80M
<pre>class control con</pre>	4			256M
Image: Solution of the second state sec				
Y # 4 Stable of disk. 40018 → 80018 Y # 4 Stable of disk. 40018 → 80018 Y # 4 Stable of disk. 1 → 7 Y # 4 Stable of disk. 1 → 7 Y # 4 Stable of disk. 90018 Stable of disk. 90018 90018 Y # 4 Stable of disk. 90018 Y # 4 Stable of disk. 90018 Stable of disk. 90018 90018 Y # 4 Stable of disk. 90018	- Gipartel data analysislenample, partel		. 10 ((JU N 453 11/12/
<pre>v Question of data and performance for the second data and pe</pre>	Ciganel data analysistexample pareli Data Graphics Statistics Uter		- 20 ((42) NU 453 M 11/12/20
<pre>ver use very very very very very very very ver</pre>	Spanet deta analysis/exemple_panet a Graphics Statistics User		• 2 0 ((3) NUM 453 PM 11/12/20
<pre>http://www.international.com/pictures/pictu</pre>	nei data analysistemarrole paneli Gaphia Sasasa User - L - X - 2 - 2 - 2 Funder	© [] [2] [] Woto wep ○ ○ of dat. 4028 → 80396	• 2 0 ((J2 NUM 453 PM 11/12/202
<pre>wrwin.</pre>	Image: State analysis iteration Image: State iteration <	© [] [2] [] Wote Hep ○ 0 of obs. <u>4025 → 10314</u> of obs. <u>4025 → 10314</u>	 2 O (Variables Ther work 	(12) NUM 153 PM 11/12/202
regen. prostatisjij gradsteijij - gradsteijij - prostatisjij - prostatisji - prostatisji - prostatisjij - pros	Image: Constraint of the second se	Image: Second	 2 • 1 Variables The variables Name 	CAP NUM 453 PM 11/12/202 F Is here Label
nviol. side satisfies in a set of the satisfies in the set of th	(data analysi learningie, ganeril inspire: Santare: User d	Image: Second	Variables	(32 NUM 453 PM 11/12/202 11/12/202 (5 here Label 11/051 (2005) or L
Partienticeticsi balinsticeticsi >> balinsticeticsi 1001. 1002. >> 1000. >> 1000. .clear .distribution .distribution .clear .distribution .distribution .idi Statio Distribution	ef data antipotexempte garell Scapho, Statuto, Uter tal	Image: Second	Yanablas Vanablas Vanablas Vanablas Suttry ceteReligon	CAP NUM 5 No 455 PM 11/12/02 - 0
<pre>NUID. List variable: eiter even (1 para) dat anitytiv@_log_para) dat: (uses mans brevippens Durvy Red (list, 100-101, 2005, 201-2012, Rousehold) - isid STATUS DISTS FAUS KEID KKID KKIPITTOD - isid di STATUS DISTS FAUS KEID KKIPITTOD - isid di STATUS - use % ipani dat anitytivangi gasal dat: (total mans brevippens Durvy Red (lists, 100-2012, Rousehold) - isid di STATUS DISTS FAUS KEID KKIPITTOD - isid di STATUS - use % ipani dat anitytivangi gasal dat: (total mans brevippens Durvy Red (lists, 100-2012, Rousehold) - isid di STATUS - use % ipani data anitytivangi gasal dat: (total mans brevippens Durvy Red (lists, 100-2012, Rousehold) - isid - total - total -</pre>	el data analysiowanyje gandi Gopher, Stoner, Der dat – Competence analysis re Varianski store re Varianski store varianski store re Varianski store varianski store v	Contemporation Contempo	Vanables	(22 NUM 53 MI 11/12/202 (
<pre>ciser swint. ciser swint. ciser swint. cises mann breelghend birry farel (100, 100-101, 2005, 201-2022, Rouehold) init STATED DISTO FAUD KED KEREITED DATE seps id * group (STATED DISTO FAUD KED KEREITED) init id BOTT ciser ciser</pre>	et data sentypidexempter generil Gosphers: Statings: User dat	Image: Second	Vanishes Vanishes Vanishes Vanishes Shifty cantelogue performer performer	CAP NUM STATUS CASTAN CASTA
<pre>- clear - clear - clear - clear - clear - clear = 0 (pace) data analysisting (pace) data - side Stantic Distric Faulty Raid REFLICED SAVET - side id Stantic Distric Faulty Raid REFLICED SAVET - side id Stantic - side id Stantic - side is the side of the</pre>	el data analysionente guerral capine Sometra una tal de analysionente una T 2 x Number T 2 x Number T 2 x Number Numb	Image: Second	Vanides Vanides Vanides Vanides Surr Surr Surr Surr surfersor protester protester protester protester	CAP NUM 433 PM 11/12/2020 1
aver 0 (pace) disk analysive (pace) dis: (India Rama Development Survey Face) (ING), FRID-101, 2005, 501-2012, Rouehold) - sees die group (STATED DISTD FAUD NED MERFLITED) - sees die group (STATED DISTD FAUD NED MERFLITED) - side die Rama Development Survey Face) (ING), FRID-101, 2005, 501-2012, Rouehold) - sees die group (STATED DISTD FAUD NED MERFLITED) - side die Rama Development Survey Face) (ING), FRID-101, 2005, 501-2012, Rouehold) - siew - siew - group (STATED DISTD FAUD NED MERFLITED) - siew - siew - siew - distr - distributive (State Staty) - siew - distributive (State Staty) - siew - siew - siew - group (State State Staty) - site - si	el das antipotransport general Gosphen Statines Uner tal	Image: Second	Variables Variables	C4P NUM C43 PM 11/12/002 C 0 C 0 C 0 C 0 C 0 C 0 C 0 C 0
Ittelia muan brevignent Survey Fael (1000, 1002-101, 2003, 2011-2012, Rouehold) - inid SIANID DISTE FAUD NET METATIO NAVIT - epen id + group (ISTERD DISTE FAUD NETA METATIO) - inid SIANID DISTE FAUD NETA METATIO) - inid Si SANTO - clear - clear - title i Botty famil dist. title is BANTO - clear - clear - title i Botty is the METATIO (1000, 1002-101, 2003, 2011-2012, Rouehold) - stati - stati - peni veriable: - inter - group (ISTERD DISTE FAUD NATE - olde:	e data analysologung genetic Sospher Stantic Vier al Arityon, Provide Stantic Vier T a X Namber T a X Namber J vazia analyson, D PSUD , clear	Image:	Vaulds Vaulds SURV SURV gentrege perform perform gentrege gentrege gentrege	CAP NUM CASING CASIN
<pre>isid STATED DISTED FAUD KED KERFLITED SAVET eqs. id e grang (STATED DISTED FAUD KED KEFFLITED) isid id STATE clear isid id STATE clear isid id StATE isid isid id StATE isid isid id StATE isid i</pre>	e das antychonope generi Gophe, Statines, Uer th	Vote trag for the subject of the subject	Vasides V	CAP NUM CAS PAIN CAS PAIN LULIZON T ILULZON T T
<pre>inid TRATED DITTO FAUL RATE REPAIR TO DAVET equal is group (ITATED DITTO FAUL RATE REPAIR TO DAVET isid is STOTE cites cites part variable: is (Iteracy) shaced time variable: is (It</pre>	e data anyonyonyong denti Graphic Stanta User al	Image: Second	Vaidds Vaidds	CAP NUM CASING CASINING CASING CASING CASING CASING CASING CASING CASING CAS
<pre>- eqs. id * group (STATED 2 SIGTE FUED REID REFLITED) - isid id BOWT - isid id BOWT - clear - use * 0 (pac) data analysis/except_pace) dat- (Edd in mach Development Survey Facei (1000, 1000-101), 2005, 2011-2012, Household) - stati - paci veriable: id (strengly blainced) - stati - ofice i sunt - ofice -</pre>	el de antonicamo de general Graphia Stantia Uer (T à X T à X T à X Stantia Uer (T à X Stantia Uer (T à X Stantia Uer (Stantia	Voto kry O	Vandes Vandes	CAP NUM CASIMA CASIMA ILIZIZON I
<pre>- equa id * group (STATIE DISTITUTE NUTE NUTE NUTE NUTE NUTE NUTE NUTE</pre>	A second se	Image: Second	Vuidds Vuidds	KARANA KANANA KANANANA KANANANANANA KANANANANANANANANANANANANANANANANANA
<pre>isid id H399TT .clear .cl</pre>	Image: State of the s	Image: Section of the section of th	Vauldes V	
ind d SWMT . clear we 'V best data salytifkreepin part fdr' Indie Start and preispent Survey Fact (1000, 1001-101, 2001, 2011-2012, Rousehold) . test part variable: MOTAT, 1 to 2 doita: 1 wait		Image:	Vaidds Vaidds	Color Note C
<pre>- clear . clear . um '% ipani data ansijrisi/stangle ganal data (india ansijrisi/stangle ganal data (india ansijrisi/stangle ganal data (india ansi Developent Survey Faet (IEGS, IEGS-IE), 2005, 2011-2012, Rouebold) . total part intibit: id introduct ansigned (india i ansit</pre>	Image: State of the s		Vanides	Constraints Constrain
<pre>. we '0 ipsoil dat aniyritinampir psoil dar [India mana Development Survey Famil (INDS, IND-101, 2001-2012, Household) . state gent variable: MOTAT, 1 to 2 doils: 1 wait command</pre>	Conservation Conservation<	Vore Kry O	Vanishes Vanishes	Kore
Lever 5 (speci data anitysiyangi pani da: (Tolisi much revigent survey funct (1600, 1801-11), 1005, 1011-2012, Household) ttet ippel variable: 500000, 15 to 2 delta: 1 unit 	4 0 0 0 4 0 0 0 0 2 0 0 0 0 2 0 0 0 0 2 0 0 0 0 2 0 0 0 0 2 0 0 0 0 2 0 0 0 0 2 0 0 0 0 2 0 0 0 0 2 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	Image:	2 Venales Venale	CAP NOM 451000 1112200 1112200 1112000 1112000 1112000 1112000 1112000 111200 111200 111
Linear mann versignen survy Feel (1800, 1807-11), 2007, 2011-2012, Mosekold) 	Consequences and a second	Image:	Vusible	CAP NUM CAP
<pre>. titl: perior variable: 16 (iteresty balanced) Line variable: 507077, 1 to 2 doils: 1 with Command 1 Command 1</pre>	Image: Second processing and the second procesing and the second processing and the second proces	Image:	Voulds V	Color Note
paral withble: if directly plained ties withble: Softration in the 2 delts: 1 unit 	Image: State Image: State<		Vusible Vusible Vusible Vusible Vosible Soft Operates Vusible Depertes Vusible Inser Top To	CAP NOM CAP NOM CAPON CAP NOM CAP N
Line variable: SUMTE: 1 to 2 delta: 1 uolt Command al Command al C	Image: Section Control	Image:	2	(2) F (M) (1) (2) (2) (2) (2) (2) (2) (2) (2) (2) (2
Command Comman	Contract	Vore verg O	Vandels Vandels Vandels Vandels Vandels Vandels Vandels Vandel Vandel	(2) * 00
Command I South So	A saviyot. A	Image:	• 2 ● 1 • 2 ● 1 • 2 ● 1	(2) P AN (2) A (2)
Command 1 1004 1004 0004 1004 <td< td=""><td>A sampair and a sampair a</td><td>Vote Mark Vote Mark</td><td>Variables Variables V</td><td>(1) (1) (1) (1) (1) (1) (1) (1) (1) (1)</td></td<>	A sampair and a sampair a	Vote Mark Vote Mark	Variables V	(1) (1) (1) (1) (1) (1) (1) (1) (1) (1)
Communit Operation MSM Minuty State 2.94 Minuty 6.84 State State 2.94 State	Constant of the second se	Image:	Vaides Vaide	(2) F M (1) F (2)
Common 5 2.794 Menuny 640 564 564 Somethy of SUBSY (3.938) (3.938) (3.938)	A sampadamong general Cogne Statute Cogne Statute Vaic	Vore verg V	Variables V	(2) * M (2) * * * * * * * * * * * * * * * * * * *
Minnung MM Sonted hy di SAU (J.2. NAM	A service of the	Vore way	► 2 ● 4 Values Values Varialis Suffreque professor Suffreque pr	C. D. K. M. STATU STATUS STATUS STAT
Solited by ud SUBNY (23 NUM	And A and A share of the second statements of the second statement	Voir Mark Voir Mark	2	(2) P (M) (1) (1) (2) (2) (2) (2) (2) (2) (2) (2) (2) (2
(A) NUM	And Sate analysis service grant of a service service of a service service of a service service of a serv	<pre></pre>	 2 ● 1 Vaules Torrers Settingo profession profes	(2) № M
	And State and Andream Regions Andream	Image:	2	(1) P

And by using household long panel we will also show it here, like we will use a household long panel data we have already defined. We will clear it here first. Then we will use a household long panel data. So, this is getting opened. In between we will show it. It is 80,036 observations. We have opened this Stata for our explanation. So, we will define a group id, group variable with the name id with egen command. So, egen, that is ID and group. So, ID I think we already. So, we have defined.

First of all, we need to understand whether they are uniquely identifying or not. So, these are there. Just a minute. What I will do, we will take it to here. And so then come back to this. And

id variable has been generated. So, this id variable is going to be your uniquely identified variable that composed of five indicators.

So, you can also take ISID with id only so that also defined to be id going to be uniquely identified. Since it is composed of time as well, so we have not taken. So, survey was not included in the group. Both the commands are followed by cross sectional. I just wanted to mention that both the information, both the commands are followed by cross sectional and time series variables in order.

We will use once again the example panel data and we will also check it xtset, whether that is perfectly with panel format or not. So, we will use the example panel data for explanation. So we will clear it and we will use the example panel data. If you simply go by xtset or tsset, so this will give us the information. This clearly confirms that, it is a strongly balanced panel data and it has the survey content from first period to second period. There are two time periods with a change of time to be one.

So, these, for these commands id variables would be numeric. We must be having id variables to be numerical. Usually, id variables are in string, but here xtset does not read the string variables.

(Refer Slide Time: 18:19)



In this case id represents the panel id information and survey represents the time component and it is a strongly balanced panel data. Ideally, we want to have a balanced dataset in our analysis, but this is not always the case. However, we can still run the model without a balance data as well.

(Refer Slide Time: 19:15)

SPE	CIALIZED PANEL COMMANDS TO UNDERSTAND DATASET	THE	
🗖 xtde	<u>s</u> cribe		
Explair	is the extent to which panel is unbalanced.		
	No. of observation)	
ld range	. ståsseribe	-	
T_i explains 100% of the household s are	1.1.2,, 40018 n = 40018 SUMMIN: 1.2,, 2 T Delte(SUMMIN: = 2 periods No. of time period (1.4) WUNTI uniquely identifies each observation Distribution of T_1: min \$4 254 504 755 555 max 2 2 2 Freq. Percent Oum. Patternet	bn	
observed	40018 100.00 100.00 11 that year, here in both the year same	no.	
periods.	40018 100.00 XX OTHOUSENDUS WERE OUSERVED.		
	6	16	
Stata/SE 15.1 - Giganel data analysidexo File Edit Data Graphics Statist	ngependa a liter Weden Help Mar D. D.	- 0	X
Review Tax		* Variables	۹ ۲
Filter commands here	1 1010 10 SUMME1	Filter variables here	
1 use 'G\panel data analysis\	. used	SURVEY IHDS1 (2005)) or L
2 reshape long casteReligion	. use "G'ipanel Gata analysis(example_panel.dta" (India Human Development Survey Panel (IHDS, IHDS-11), 2005, 2011-2012, Household)	casteReligion HQ3 1.13-15 0	Caste
4 use 'G\panel data analysis\	. xteet	paidWorker HQ14 8.4c But	15151
 isid STATEID DISTID PSUID egen id = group(STATEID 	panel variable: id (strongly balanced) time variable: SURVEY, 1 to 2	id group(STATE)	
7 isid id SURVEY	delta: 1 unit	IGVA	
8 clear 9 use "G\panel data analysis\	. xtdesc		
10 xtset	r(199);		
12 stdes	. xtdes		
	id: 1, 2,, 40018 n = 40018		
	SUMVET: 1, Z,, Z T = Z Delta(SUMVET) = 1 unit	Properties	÷
	<pre>Span(SURVEY) = 2 periods (id*SURVEY uniquely identifies each observation)</pre>	 Variables 	
	Distribution of T 1: min 5% 25% 50% 75% 95% max	Label IHDS1 (200	J05) or №
	2 2 2 2 2 2 2 2	Expense int	
	Freq. Percent Cun. Pattern	Value label SURVEY	
	40018 100.00 100.00 11	A Data	
	40018 100.00 xx	Filename example_pu	panel dti
		Label India Huma Notes	
		Variables 7 Observations 80.034	
	Command	Size 2.29M	
		Memory 64M Sorted by id SURVEY	c.
G\Stata15		CAP N	KUM OI
	D 🗿 🗀 👩 🛄	- 10 4 to N 454	PM

The usual way of explanation of almost all the dataset we do with their description. Here we do xtdescribe it simply gives us the explanation xtdes. It gives us explanation for how many observation 40,018 observations are there in total in two time periods. There t is two, over a span

of two periods is given. It also gives information of the pattern. Pattern suggests that it is one stands for one period information means one observation in that year and here in both the years same number of households are observed. So, that is why in both, here it is one and one is given. If same number of house information are not available then there must be a dot or not available indicators must have been given. I think I have already explained. I must proceed.

(Refer Slide Time: 19:45)



Coming to summarize the way we summarize, we generally summarize in the cross section, between analysis, basically, between the observation we compare in case of summarize. But in

panel data we have all the information basically, between the observation and within the time, within the household between the time period that is called within analysis is also important and overall explanation is also possible. So, xtsum if you do it, will find out the information of all that is being discussed here.

So, everything is mentioned. For your explanation I just wanted to give that overall description, like ordinary statistics that are based on total number of observation is given in overall indicators. Between is important so far as calculation of the summary statistics of number of households, regardless of time period. Basically, it is cross sectional information given so far as between is concerned, because time component is not considered while understanding their summary statistics.

(Refer Slide Time: 20:46)

	of number of households regardless of time period. Within: summary statistics of two time periods regardless of number of households.	er
	In the above result, gross receipt is more between household than it changes over time for the same household.	S
	N is total number of observation while n is observation of each entity.	n
	Within variation of id variable is zero. Since id does not var over time. Similarly, survey variable has between variation zero as time does not vary over id.	y n
🗿 - swa	yan 🧕	18

□ <u>xtsum</u> marize						
It gives summary statistics. It	Variable	Mean	Std. Dev.	Min	Max	Observations
also separates within (over	SURVEY overall	1.5	. 5000031	1	2	N = 80036
	within		. 5000031	1	2	T = 2
time) and between (over individuals) variation.	casteR~m overall between within	3.631939	1.367554 1.325354 .3369073	2 2 1.131939	7 7 6.131939	N = 00014 n = 40010 T = 1.99945
🗖 llava stata lista threa	grossR-s overall between	214991.7	454461	0	1.84e+07 1.36e+07	n = 11731
🖬 Here, stata lists three	within		300125.6	-8965008	9394992	T-bar = 1.32501
different types of statistics.	paidWo~r overall between	14259.39	63311.48 54255.16	0	1600000	N = 14976 n = 11309
uncrent types of statistics.	within		31822.74	-785740.6	814259.4	T-bar = 1.31495
Overall: ordinary statistics	busine-n overall between within	1.901107	.7280784 .7107795 .2751249	1 1 .9011075	3 3 2.901107	N = 16614 n = 12200 T-bar = 1.35293
that are based on total	id overall	20009.5	11552.27	1	40018	N = 00036
number of observation.	vithin		11552.35	20009.5	40018	T = 2
Between: calculated on the basis of summary statistics of	16VA overall between within	11.35801	1.400061 1.368424 .463737	1.149598 1.149598 8.546378	16.72569 16.72569 14.16965	N = 15407 n = 11633 T-bar = 1.32442
🕘 saayahi 🧕						17

Within is given concerning the time period, regardless the number of households. So, in the above result, gross receipt is more between households than it changes over time for the same period, gross receipt which is shown here. between is much higher. Between basically across the household, the gross receipt is much higher as compared to the within. So, the N is total number of observation, while small n is observation on each entity. While understanding small n and capital N is giving, the right hand side and within variation of id variable is 0 and since id does not vary over time. Id remains same over time. So, the variation within id and their variation is 0. Their variation basically the standard deviation is 0. Coming to the survey variable has between variation 0 as time does not vary over time, over id.

(Refer Slide Time: 22:18)



In all our variables between variation is more than the within variation. You please mark everywhere that we have already mentioned. This means these variables vary between households. Between households there are much variation, but the same household over time the variation is compared to lesser. So, this is one of the findings of the data we have so far summarized.

(Refer Slide Time: 22:51)

PANEL MODELS IN STATA	
Pooled OLS Model: This is done by simply running a regression model of dependent and independent variables. reg IGVA businessLocation paidWorker casteReligion Or reg IGVA i.businessLocation paidWorker i.casteReligion	the
🎯 - swayani 👲	20

Review T 4	×	within			0 20009.	5 2000	9.5	T = 2			 Variab 	les	T 4 >
 Filter commands here 	0 .							N - 15407			1		
# Command in		betwee	1 11.330	1.3684	24 1.14959	8 16.72	569	n = 11633			Na	ne	Label
use 'G\panel data analysis\		within		.4637	37 8.54637	8 14.16	965 T-	bar = 1.32442			SUR	RVEY .	IHDS1 (2005) or L.
2 reshape long casteReligion		an long i bu									cast	teReligion	HQ3 1.13-15 Caste
3 clear		reg lova 1.bu	SINESSLOCAT	ion paidwor	ter 1.castes	eligion					gro	ssReceipts	HQ14 8.3 Busns1:
use 'G\panel data analysis\		Source	55	df	MS	Nunber	of obs	- 14,392			paid	dWorker	HQ14 8.4c Busns1
isid STATEID DISTID PSUID	-					F(0, 14	383)	= 598.12			bus	inessLoca	HQ14 8.7 Busns1:
5 egen id = group(STATEID		Model Residual	6933.5966		866.699574	Prob >	E	- 0.0000			id		group(STATEID DL.
7 isid id SURVEY	_	Residual	20011.1900	14,303	1.11903073	Adi R-s	cu guared	- 0.2492			IGV.	A	
8 clear		Total	27775.0921	14,391	1.93003211	Root MS	3	- 1.2038					
a use 'G.\panel data analysis\													
10 xtset	. –		1.034	Conf	and free		PS I F I	1953	Internall				
11 HOEK 19	· _		1010	coer.	303. 511.		EXIT.	1999 Cont	. incervar)				
11 storm		busines	sLocation										
		Other	fixed 2	7329866	.0234135	31.31	0.000	6870931	7788801				
14 reg IGVA i businessLocation.		ouner											
14 reg IGVA i.businessLocation		other	Moving 3	.169172	.0280563	6.03	0.000	1141779	.224166		Proper	rties	Q :
14 reg IGVA i.businessLocation.		ocher P	Moving 3 aidWorker	.169172 6.24e-06	.0280563	6.03 39.55	0.000	.1141779 5.93e-06	.224166 6.55e-06		Proper	rties •	¢ :
14 reg IGVA ibusinessLocation.		p	Moving 3 aidWorker	.169172 6.24e-06	.0280563 1.58e-07	6.03 39.55	0.000	.1141779 5.93e-06	.224166 6.55e-06		Proper Vari	eties ables	9
14 reg IGVA ibusinessLocation.		p cast	Moving 3 widWorker eReligion	.169172 6.24e-06	.0280563 1.58e-07	6.03 39.55	0.000	.1141779 5.93e-06	.224166 6.55e-06		Proper • • Vari Nam	ables	SURVEY
14 reg IGVA i businessLocation.		p. cast	Moving 3 aidWorker eReligion OBC 3 Dalit 4	.169172 6.24e-06 4693355 6807797	.0280563 1.58e-07 .0262581 .0336599	6.03 39.55 -17.87 -20.23	0.000	.1141779 5.93e-06 5208048 - 7467575	.224166 6.55e-06 4178662 - 614802		Proper Vari Narr Labr	ities ables ic ic	SURVEY DIDS1 (2005) or Di rot
14 reg IGVA i.business.location.		p cast	Moving 3 aidWorker eReligion OBC 3 Dalit 4 divasi 5	.169172 6.24e-06 4693355 6807797 -1.493521	.0280563 1.58e-07 .0262581 .0336599 .0484026	6.03 39.55 -17.87 -20.23 -30.86	0.000	.1141779 5.93e-06 5208048 7467575 -1.588397	.224166 6.55e-06 4178662 614802 -1.398646		Proper	rties ables el el en rat	4 : SURVEY 0-IDS1 (2005) or 0-1 int 57.04
14 reg IGVA i.businessLocation.		cast N	Moving 3 aidWorker eReligion OBC 3 Dalit 4 divasi 5 Muslim 6	.169172 6.24e-06 4693355 6807797 -1.493521 [2.3550901	.0280563 1.58e-07 .0262581 .0336599 .0484026 .0323887	6.03 39.55 -17.87 -20.23 -30.86 -10.96	0.000	.1141779 5.93e-06 5208048 7467575 -1.588397 4185762	.224166 6.55e-06 4178662 614802 -1.398646 291604		Proper	rties ables he e r nat se label	SURVEY BHDS1 (2005) or BH Int S7.01 SURVEY
14 reg IQVA i.businessLocation	0	cast ristian, Sikh	Moving 3 waidWorker eReligion OBC 3 Dalit 4 divasi 5 Muslim 6 i, Jain 7	.169172 6.24e-06 4693355 6807797 -1.493521 D.3550901 .0930179	.0280563 1.58e-07 .0262581 .0336599 .0484026 .0323887 .0600656	6.03 39.55 -17.87 -20.23 -30.86 -10.96 1.55	0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.121	.1141779 5.93e-06 5208048 7467575 -1.588397 4185762 0247184	.224166 6.55e-06 4178662 614802 -1.398646 291604 .2107542		Proper * * Vari Labi	rties ables el el ruit ie label es	G : SURVEY BHDS1 (2005) or BH INT SURVEY
14 reg IQVA i.businessLocation	c	cast cast listian, Sikh	Moving 3 aidWorker eReligion OBC 3 Dalit 4 divasi 5 Muslim 6 , Jain 7	.169172 6.24e-06 4693355 6807797 -1.493521 D:3550901 .0930179	.0280563 1.58e-07 .0262581 .0336599 .0484026 .0323887 .0600656 0257891	6.03 39.55 -17.87 -20.23 -30.86 -10.96 1.55 439.16	0.000 0.000 0.000 0.000 0.000 0.000 0.121	.1141779 5.93e-06 5208048 7467575 -1.588397 4185762 0247184 11.2751	.224166 6.55e-06 4178662 614802 -1.338646 291604 .2107542 11.3762		Proper Proper Value Nam Labe Form Value Note Value	rties ables be el el el bel el bel es a	 SURVEY H-DS1 (2005) or BH INT INT INT SURVEY
(4 reg IQVA i burnentocation.	c	cast ristian, Sikh	Moving 3 aidWorker eReligion OBC 3 Dalit 4 divasi 5 Muslim 6 , Jain 7 cons	.169172 6.24e-06 4693355 6807797 -1.493521 D:3550901 .0930179 11.32565	.0280563 1.58e-07 .0262581 .0336599 .0484026 .0323887 .0600656 .0257891	6.03 39.55 -17.87 -20.23 -30.86 -10.96 1.55 439.16	0.000 0.000 0.000 0.000 0.000 0.000 0.121 0.000	.1141779 5.93e-06 5208048 7467575 -1.588397 4185762 0247184 11.2751	.224166 6.55e-06 4178662 614802 -1.398646 291604 .2107542 11.3762		Proper • Vari Labi Type Form Valu Note Note - Data	rties ables he el ruit ie label es a aume	g SURVEY BHDS1 (2005) or BH Int SURVEY example_panel dta
14 reg IQVA ibunentication.	8	cast p cast iristian, Sikh	Mowing 3 aidWorker eReligion OBC 3 Dalit 4 divasi 5 Muslim 6 , Jain 7 cons	.169172 6.24e-06 4693355 6807797 -1.493521 D-3550901 .0930179 11.32565	.0280563 1.58e-07 .0262581 .0336599 .0484026 .0323887 .0600656 .0257891	6.03 39.55 -17.87 -20.23 -30.86 -10.96 1.55 439.16	0.000 0.000 0.000 0.000 0.000 0.000 0.121 0.000	.1141779 5.93e-06 5208048 7467575 -1.588397 4185762 0247184 11.2751	.224166 6.55e-06 4178662 614802 -1.398646 291604 .2107542 11.3762		Proper Vari Labe System Valu Valu Form Valu Valu	rties ables be el rat ie label es a aume el	SURVEY HIDST (2005) or HI Int %7.0F SURVEY example_panel dta India Human Deve
14 reg IQVA ibunentication.	8	p cast k ristian, Sikh	Moving 3 aidWorker eReligion OBC 3 Dalit 4 divasi 5 Muslim 6 , Jain 7 cons	.169172 6.24e-06 - 4693355 - 6807787 -1.493521 D-355990 .0930179 11.32565	.0280563 1.58e-07 .0262581 .0336599 .0484026 .0323887 .0600656 .0257891	6.03 39.55 -17.87 -20.23 -30.86 -10.96 1.55 439.16	0.000 0.000 0.000 0.000 0.000 0.000 0.121 0.000	.1141779 5.93+-06 5208048 7467575 -1.588397 4185762 0247184 11.2751	.224166 6.55*06 4178662 614802 -1.338646 .231604 .2107542 11.3762		Proper Vari Labe Labe Labe Valu Valu Labe Valu	rties ables be el e tabel es aume el es es es es es es es es es es	g SURVEY SURVEY int %7.06 SURVEY example_panel dt/ India Human Deve
14 reg IQVA: Ibunentication.	c:	Cast P Cast In Infistian, Sikh	Noving 3 aidWorker eReligion OBC 3 Dalit 4 divasi 5 Muslin 6 Muslin 7 cons	.169172 6.24e-06 4693355 6807797 -1.493521 D≥.355090 .0930179 11.32565	.0280563 1.58e-07 .0262581 .036559 .0484026 .0323857 .0660656 .0257891	6.03 39.55 -17.87 -20.23 -30.86 -10.96 1.55 439.16	0.000 0.000 0.000 0.000 0.000 0.000 0.121 0.000	.1141779 5.93e-06 5208048 7467575 -1.588397 4185762 0247184 11.2751	.224166 6.55e-06 4178662 614802 -1.39864 .291604 .2107542 11.3762		Proper Varia Varia Labo Form Value Val	rties ables be el e nut ie label es aume el es ables ables antonos	SURVEY HOS1 (2005) or HI int ST66 SURVEY example_panel dta India Human Deve J Bons6
14 reg IQVA:burnent.oration.	c:	cast P cast iristian, Sikh	Moving 3 aidWorker eReligion OBC 3 Dalit 4 divasi 5 Muslin 6 , Jain 7 Cons	.169172 6.24e-06 4693355 6807797 -1.493521 D≥.355090 .0930179 11.32565	.0280563 1.58e-07 .0262581 .0336559 .0484026 .032887 .0600656 .0257891	6.03 39.55 -17.87 -20.23 -30.86 -10.96 1.55 439.16	0.000 0.000 0.000 0.000 0.000 0.121 0.000	.1141779 5.93e-06 - 5208048 - 7467575 -1.588397 - 0247184 11.2751	.224166 6.55e-06 4178662 614802 -1.398646 291604 .2107542 11.3762		Proper Proper Value	rties ables as e e label e label es ables esvabors	SURVEY HOSS (2005) or H in SURVEY example_panel dta losta Human Deve / B0036 2.37M
M ng GVA bannestoratos.	c	venand	Noving 3 aidMorker eReligion OBC 3 Dalit 4 Dalit 4 Muslin 6 Jain 7 cons	.169172 6.24e-06 4693355 6807797 -1.493521 ▷.3550901 .0930179 11.32565	.0280563 1.58e-07 .0262581 .0336559 .0484026 .032887 .0600656 .0257891	6.03 39.55 -17.87 -20.23 -30.86 -10.96 1.55 439.16	0.000 0.000 0.000 0.000 0.000 0.121 0.000	.1141779 5.93e-06 5208048 7467575 -1.985397 4385395 0247184 11.2751	.224166 6.55e-06 4178662 614802 -1.398646 .291654 2107542 11.3762		Proper Proper Vari Lob Form Proper Lob Proper	rties ables se el el ables el el el el el el el el el el el el el	9 50/8949 9+D51 (2005) or III INT 51/20 9-2
M ng GVA banestorios.	¢	vinit p cast h ristian, Sikh remand	Moving 3 aidWorker eReligion OBC 3 Dalit 4 divasi 5 Muslin 6 Muslin 7 cons	.169172 6.24e-06 4693355 6807797 -1.493521 2.3550901 .0930179 11.32565	.0280563 1.58e-07 .0262581 .0336599 .0484026 .02848026 .0600556 .0257891	6.03 39.55 -17.87 -20.23 -30.86 -10.96 1.55 439.16	0.000	.1141779 5.53e-06 5208048 7467575 15855762 4185762 0247184 11.2751	.224166 6.55e-06 4178662 614802 -1.398646 291604 .2107542 11.3762		Proper Property And	rties ables he el s nut ie label es ables ervations ervations nory ed by	3 3

Coming to the panel models in Stata, we are now going to discuss three important aspects of analysis. One is through pooled panel, pooled OLS model that is simply the appending approach we discussed, but we are, in this panel only will simply give the ordinary regression technique. So, we will mention here and we will also operate it and we will take this command and that will be derived. We will operate and it is here.

So the result is there. This is a simple OLS regression result, but our data is pooled type. It is in panel format, but it has only considered a pooled format information by considering entire observations. So, it has given the result accordingly.

(Refer Slide Time: 23:51)

Source Model (Residual 20 Total 21	SS 5933.5966 0841.4955 7775.0923	df 5 8 866. 5 14383 1.44 1 14391 1.93	HS 699574 903675 003211	Num F(Pro R-s Adj Roo	ber of ob 8, 14383 b > F quared R-square t MSE	s = 14392 = 598.12 = 0.0000 = 0.2496 d = 0.2492 = 1.2038		The coefficients are treated as usual regression
businessL Other fi Nov paic casteR D Adi Nu Christian, Sikh, S	1GVA ocation ixed 2 ring 3 Worker OBC 3 alit 4 rasi 5 alit 4 Fain 7	Coef. .7329866 .169172 6.24e-06 4693355 6807797 -1.493521 3550901 .0930179	Std. Err. .0234135 .0280563 1.58e-07 .0262581 .036559 .0484026 .0323887 .0600656	t 31.31 6.03 39.55 -17.87 -20.23 -30.86 -10.96 1.55	<pre>₽> t 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.121</pre>	<pre>[55% Conf. .6870931 .1141779 5.93e-06 5208048 7467575 -1.588397 4185762 0247184</pre>	Interval] .7788801 .224166 6.55e-06 4178662 614802 -1.398646 291604 .2107542	models. Example: as compare to home based business other fixed premises and moving business perform better.
	_cons	11.32565	.0257891	439.16	0.000	11.2751	11.3762	

We will compare this. We will discuss this. The coefficients are treated as usual regression model. Example, as compared to home based business other fixed premises and moving business perform better that we discussed earlier.

(Refer Slide Time: 24:10)



922 548 535 10%A Coef. cation ing 3 .1670682 Norker 4.00e-06	Obs pe F(8,33 Prob > Std. Err. .0457556 6 .0587889 2	group: nin = avg = max = 1) = F = t P> t 53 0.000 84 0.005	1 1.3 2 42.18 0.0000 [95% Conf.	. Interval]	 The random Value Nune SUPEr Control of the random Value optimizing optimizing optize optimizing optimizing optimize optimizing optize optimizing o	bits here Label BHD51 (2005) or Li n HQ3113-15 Caste st HQ14.83 BusesL HQ14.84C BusesL HQ14.84C BusesL HQ14.84C BusesL HQ14.84C BusesL group(STATEID DL Group(STATEID DL)
922 548 536 10VM Coef. cation xed 2 .2999303 1.670682 Mocker 4.00e-06	F(8,33 Prob > Std. Err. .0457556 6 .0587889 2	nin = avg = nax = 1) = F = t P> t 53 0.000	1 1.3 2 42.18 0.0000 [95% Conf.	. interval]	Name SURTY castReligio grootRepp padWorke bournet.loc id IGVA	Label BHDS1 (2005) or L. n HQ3 113-15 Caste s HQ14 8.3 Buses1 HQ14 8.4 Buses1 n HQ14 8.7 Buses1 group(STATEID DL.
546 536 .1933 10VA Coef. .cation .eation .eation 	F(8,33 Prob > Std. Err. .0457556 6 .0587889 2	avg = nax = 1) = F = t P> t 53 0.000 84 0.005	1.3 2 42.18 0.0000 [95% Conf. .2092183	. Interval]	SURVEY casteReigo grosReego p.gadNober businensLoc id IGVA	BHDS1 (2005) or L. in HQ3 1.13-15 Caste. is HQ14 8.3 Busns1: HQ14 8.4c Busns1: group(STATEID DL.
336 16WA Coef. cation .2989303 img 3 .1670682 Morker 4.00e-06	F(8,33 Prob > Std. Err. .0457556 6 .0587889 2	nax = f = = t => t 53 0.000 84 0.005	2 42.18 0.0000 [95% Conf. .2092183	. Interval]	casteReligio grossRecep pardWooker business.co id IOVA	n HQ31.13-15 Caste. ns HQ14.8.3 Busns1: HQ14.8.4c Busns1: group(STATEID DL
1993 Coef. lGVA Coef. cation	F(8,33 Prob > Std. Err. .0457556 6 .0587889 2	1) - f	42.18 0.0000 [95% Conf. .2092183	. Interval)	grossReceip pardNotker businetsLoc id IGVA	 HQ14 8.3 Busns1: HQ14 8.4: Busns1: HQ14 8.4: Busns1: group(STATEID DL)
1933 10VA Coef. cation xed 2 .2989303 .1670682 Morker 4.00e-06	Prob > Std. Err. .0457556 6 .0587889 2	E = = = = = = = = = = = = = = = = = = =	0.0000 [95% Conf. .2092183	. Interval]	paidWorker businessLoc id IGVA	HQ14.8.4c Busns1: a. HQ14.8.7 Busns1: - group(STATEID DL
lGVA Coef. cation .2989303 sing 3 .1670682 Worker 4.00e-06	Std. Err. .0457556 6 .0587889 2	t P> t 53 0.000 84 0.005	[95% Conf.	. Interval]	businessLoc id IGVA	a HQ14 8.7 Busins1: group(STATEID DL
lGVA Coef. cation .2989303 sed 2 .2989303 ing 3 .1670682 Worker 4.00e-06	Std. Err. .0457556 6 .0587889 2	t P> t 53 0.000 84 0.005	[95% Conf.	. interval]	id IGVA	group(STATEID DL
Coef. Cation xed 2 ing 3 .1670682 Norker 4.00e-06	.0457556 6 .0587889 2	53 0.000 84 0.005	.2092183	. Interval)	IGVA	
cation xed 2 .2989303 ing 3 .1670682 Worker 4.00e-06	.0457556 6 .0587889 2	53 0.000 84 0.005	.2092183			
xed 2 .2989303 ing 3 .1670682 Worker 4.00e-06	.0457556 6 .0587889 2	53 0.000 84 0.005	.2092183			
ing 3 .1670682 Worker 4.00e-06	.0587889 2	84 0.005		. 3886423		
Worker 4.00e-06			.0518021	.2823343		
1.VV0-VV	2 404-02 16	4 0.000	1 \$24.06	4.474-06		
	1.404-07 10		0.016-00	1.1/6-00		
ligion 🔓						
OBC 30842451	.0630215 -1	34 0.181	20781	.0393199	Properties	ę
11t 4126534	.1068553 -1	18 0.236	3360428	.0829749	🔒 + +	
lin 60095152	.3137962 -0	03 0.976	- 6247686	6057382	 Variables 	
ain 70711652	1510055 -0	47 0.637	3672384	.2249081	Name	SURVEY
					Label	
11.20352	.0749093 149	56 0.000	11.05664	11.35039		
iona u 1.2744191					Volum Label	
igma_e .89188341					Notes	
rho .67124485	(fraction of w	riance due t	o u_i)		- Data	
Lat. 1/11042 33311	- 2.42	Reals >	F = 0.0000		Filename	example_panel d
1-0. (11002, 3321)	- 2.07	1100 /	1 - 9.9999		Label	India Human Dei
					 Notes 	
					Variables	
		_	_		Coservations	80,036
					Size	2.37M
	111 4 - 12639 111 4 - 12639 111 6 - 005552 111 6 - 005552 111 6 - 005552 112 0055 11 20552 11 20555 11 20555 11 20555 11 20555	111 010351 10093 - 2 111 010351 10093 - 2 11500153 1315003 - 0 	111 0	<pre>111 1</pre>	11	11 1.15 1.15 0.15 0.05 0.15 0

Coming to the fixed effect model. So, the fixed effect and random effect model and which one is going to be more suitable in this panel data is going to be discussed right now. Interestingly, while entering the syntax or the command, we have to specify whether it is fixed effect with fe. And if you also wanted to understand the heteroscedasticity information or the robust information, robust an option to be added here. But let us compare, understand the fixed effect regression model through xtreg with this. Let me mention that, it is here and we will do that mentioned it. And this is all fixed effect model.

The result is in front of you. And the correlation by assumption we know that the correlation between the error term and the independent variable or the explanatory variables are non-zero. Basically in the random effect model we have mentioned that there occurs a stochastic relationship between the random, the error term and the explanatory variables. But the assumption here in the correlation, there exist certain correlation between these two. So, that is the reason why the correlation value here is having some positive number. (Refer Slide Time: 25:48)



So, coming to the errors, are correlated with the regressors in fixed effect model. In case of random, by assumption, it will be 0 and we will show it. Here other important aspect that the overall model is statistically significant as mentioned with this and coming to the sigma u and sigma e as we already pointed out in our theoretical portion of panel data that sigma u is the standard deviation of residual within the group of alpha i the constant effect, the fixed effect without the time component. And it is, the deviation, the sigma value is given. And sigma e is having the standard deviation of the residuals with the time component. That is epsilon. Both the time component is there.

The rho gives the value of 67.12 percent that explain the variation is due to differences across panel, the difference is of 67.12 percent. And there are some indicators we are highlighting here having non-significant. So, as per our model, since we did not specify the model so systematically we have simply shown the results. So, many variables are not coming out to be significant.

(Refer Slide Time: 27:21)



Coming to the random effect model, we will only add re there to our model. Random effect, we simply changed to re instead of fe. It gives the random effect results. As I already mentioned from the beginning that the correlation by assumption is 0. There occurs a stochasticity between these two variables. So, this is assumed to be 0.

Coming to the explanation once again, the interpretations are little tricky in this case since they include within entity and between entity effect. It is interpreted as the average effect of other fixed premises as compared to the home-based business is positive and home-based and the fixed

premises is positive and 0.67, I already given in the coefficient, 0.67 more on the log performance on the business, 0.67 more so far as the fixed business as compared to home business is concerned. So, that is more important and rest details you will get it from our document.

(Refer Slide Time: 28:47)

	FEM Vs. REM		
Hau is b	usman specification test gives information on which etter.	n moo	lel
quietl fe	y xtreg IGVA i.businessLocation paidWorker i.casteR	Religio	on,
estima	ate store fixed		
quietl re	y xtreg IGVA i.businessLocation paidWorker i.casteF	Religio	on,
estima	ate store random		
hausn	nan fixed random		
🙆 - Swayani	<u>@</u>		25
Stata/SE 15.1 - Gilgarnel data analysis File Edit Data Graphics Sta	იიიყნელისეს მის (სილ Workson Herg 24 - 11 - C C		- 0 - X -
Review T 3	x	Variables	τů×
Filter commands here	· WEATHER BOOK ANTONY	Name	es here
1 use 'G\panel data analysis\	. heusman fixed random	SURVEY	IHDS1 (2005) or L.
reshape long casteReligion clear use 'G\panel data analysis\ isid STATEID DISTID PSUID egen id = group(STATEID	Bote: the rank of the differenced writence matrix (7) does not equal the masher of coefficients being tested (0); be sure this is shak yoe expect, or there may be problem computing the test. Lamine the copper of your estimators for mything mempercies and possibly consist sollary your writelikes so that the coefficients are on a similar scale. — coefficients —	casteReligion grossReceipt paidWorker businessLoca id	HQ3 1.13-15 Caste HQ14 8.3 Busns1: HQ14 8.4c Busns1: HQ14 8.7 Busns1: group(STATEID DL.
7 isid id SURVEY 8 clear 9 use "G\panel data analysis\	(b) (B) (b-0) sqt(disq(V_b-V_B)) fixed random Difference S.E.	IGVA _est_fixed _est_random	esample() from es esample() from es
10 xtset 11 xtdex 199 12 xtdes 13 xtsum	2 2499303 .515458 -315199 .039227 3 216962 .215928 -027244 .555244 pal80criter 4.60e-65 .54e-64 -1.81e-65 .1.85e-67 casterbelipen 3 -0462451 -43338 .347093 .656849		
14 reg IGVA (businessLocation.) 15 xtreg IGVA (businessLocati.)	4 - 126534 - 6547333 .5281993 .1008686 5 .1640611 -1.447422 1.611483 .1693958	Properties	₽×
16 xtreg IGVA i businessLocati	6 - 0095152 - 3173973 3078821 3118227 7 - 0711653 1010872 - 1343524 1172561	4 Variables	
 quietly streg IGVA i busines estimate store fixed 	· · · · · · · · · · · · · · · · · · ·	Name	SURVEY NUDEL CONTENT IN THE
19 quietly streg IGVA i busines	b = consistent under No and Ha; obtained from xtreq B = inconsistent under Ha, efficient under Ho; obtained from xtreq	Type	int (2005) or IHD
20 estimate store random 21 hausman fixed random	Test: No: difference in coefficients not systematic	Format Value Labor	%7.01 SURVEY
-	$ch(2 \mathbf{T}) = (h_1h_1)^*((2 h_1V R))^*(-111 h_2R)$	Notes	
	= 238.25 Route () = 4440	Data Filename	example panel dta
	rrob>chl2 = 0.0000	Label	India Human Develo
		Variables	
	Command	Observations	80,036 1 SIM
		Memory.	61M
		Sorted by	id SURVEY
G\Stata15			Stor PM
🧶 🕞 🖉 🖌			11/12/2020



We will come to the comparison between the random effect and the fixed effect model. As we repeatedly mentioned earlier that we require a Hausman specification test that really identifies which model to be better. We will compare these the way we did it earlier. We will compare also the same thing. We will quietly as the command we stated earlier that we will take the quietly command every time the way we taken quietly. Then we will also estimate its store value and store it. Basically, we will estimate and store it in every cases. Every time we will go by that. Then our results will be derived very quickly.

So, we derived fixed effect. We derived random effect. Then we also check a Hausman test. So, all the results are derived in front of you. After estimating all those detail, the comparison through the Hausman test is very clearly mentioned. As I told you, fixed effect and its coefficients are mentioned, random effect and its coefficients are mentioned, as per our Hausman effect if you remember on the 36th lecture we discussed that why Hausman is required and what is the formula for it.

We take the chi square estimation of these two differences. So, we take the difference and divide it with their standard deviation. There square root of these variation. So, these variations are defined and defined to be significant. So, this significance really statistically significantly violating our null hypothesis. So, null hypothesis is that the difference in coefficients are systematic. So, there are not systematic, sorry. The difference in coefficient, these two coefficients are non-systematic. So, they are very random, but we are rejecting that null hypothesis based on the significance level. Our significance level confirms that our model is statistically significant and we are rejecting a null hypothesis. So, that means, the difference in coefficients are systematic. So, there is a very clear understanding based in the model. And based on that, we say that when they are very systematic that means the correlation between the error term and the explanatory variables are there and they are not random.

So, the assumption of the random effect is violated. So, this model, Hausman test confirms to the fact that we have to apply the fixed effect model. So, that is what we have derived. And this test confirms to the fact that the fixed effect model is more appropriate.

So, that is all about the discussion in the entire lectures of the 40 modules. This is, in fact, our 40th module and we have discussed everything about almost all basics of panel model. There are many details of panel models also. I think we have given a very good base for all of you to understand the unit level data and their structure of making a panel and how to analyse panel with the basic guidance have already been given. So, with this, I think it is the time to conclude and mention that we have systematically proceeded from the unit level data and the basics of Stata, then conceptualizing Stata, we codify Stata.

Then we enter with various components of panel and we also analyse panel. I hope you guys enjoy in between and any kind of doubts persist please do not hesitate to raise it and we will be very happy to address it. With this, I am closing the lecture and expecting your better participation and performance in the exam. Thank you.