Exchanging Data between R and Excel

CAS Spring Meeting 2012 Kirsten Singer ms.kirsten.singer@gmail.com

How-to get Excel data into R

- I. Excel data sample
- II. Import data to R from csv
- III. Copy data to R from clipboard

Appendix I. Packages

Appendix II. Exporting data from R

Excel data sample

Data features

- Spaces within the column headingsMissing values
- Mixture of character fields, numeric fields, and date fields

	A	D.	C	D	E	E	G	H
1	Polity ID	Claim ID	Pol Ett Date	Acc Date	Palid Indiemnity	Paid Expense	Case RSV Indemnity	Case Reserve Expense
2	P000001	C000001	2011-01-01	2011-06-30		160.04		207.45
3	P0000002		2011-01-02	2011-07-01	51.45	57.16	63.52	70.57
4	P0000003	0000003	2011-01-03	2011-07-02	52.4	58.22	64.69	71.67
5	P000004	C000004	2011-01-04	2011-07-03				
ь		CDOODES	2011-01-05	2011-07-04	22.52	25.02	27.8	30.89
7	P000006	C000000		2011-07-03	62.11	69.01	76.68	83.2
8	P000007	C000007	2011-01-07	2011-07-06		16.03	17.82	19.79
9	P0000008	0000008	2011-01-08	2011 07 07	66.12	73.46	81.63	90.7
10	2000009	C000009	2011-01-09	2011-07-08	168.08	186.76	207.51	230.56
11	P0000010	0000010	2011-01-10		142.83	158.7	176 33	195.92

 \checkmark Date fields have been formatted in the proper way, as yyyy-mm-dd

-	

Basic command: > claims <-read.csv (file="claims.csv") • Use a few additional specifications to ensure data is imported with correct data types and blanks are handled correctly ✓ Use colclasses to define data types, and na.string to assign blanks as NA data ! Note that R uses a default date format of yyyy-mm-dd, so you must change the way dates are displayed, in Excel, to this format • In Excel, export data to csv file and save it to any logical directory • Change working directory in R using the file menu — point to the file's location | Part | Par

Ва	sic con	nmand:	> claim	s <-read.ta	ble (file="cl	ipboard")		
•						ead.table nee	eds a couple ext k data	ra
✓							works with data	on the
✓	Head	er=TRU	E lets R kn	ow the dat	a has heade	rs (default is	FALSE for read.	table)
Ţ					to yyyy-mr			,
	cqu.	. cu uut	а ргер. со		, , , , ,			
				crevious exampl				
10	datatype:	' already	defined in ;					
) ic	datatype: opy data wad in da	' already to clipho	defined in ; and, then	previous exampl	la			
o fo	datatypes opy data wad in da singv-ree	' already to clipho ta d.teble(E	defined in ; and, then	previous exampl	la	Jalalypes, beader	-INUE, sep-*\L*)	
o fo	datatype: opy data wad in di sinst-res	' already to clipho	defined in ; and, then	previous exampl	la		-INUE, sep-*\L*)	
) (1) (1) (1) (1	datatypes opy data wad in da sing/vite brok isso sing	' already to clipho ta d.table(E rotal dava	defined in ; and, then Lit="cliphos:	previous exampl	la ,colClesses	detetypes,breder		eserve Espans
FI FI V UI V UI	datatypes upy data wad in da sing/-res brok issu- sing blicy.ID	' already to clipho ta d.table(E rotal dava	defined in ; and, then Lit="cliphos:	orevious exampled, neurotrings Acc.Date Par	la ,colClesses	detetypes,breder	-TRUE, sep-"\L") RSW. Indemnity Care. Re	
) #2) (1) (1) (1) (1	datatypes upy data wad in da sing/-res brok issu- sing blicy.ID	' already to clipto ta d.teble(f eved deve Claim.ID	defined in p and, then lie-"clipbon Pol.Eff.Dure 2011-01-01	cevious exampl d',nm.strings Acc.Date Fm: 2011-06-30	le , cultieseter. id.Indemnity Fer	deletypes,beede: id.Expense Case.i	RSV. Indemnity Care. Re	207-4
fr fr fr fv fv	datatypes opy data wad in da alast-rea alast- alast- blicy.ID possoon possoon possoon	" already to clipho ta d.teble(f roted deca Claim.ID: C000001	defined in p and, then lie-"clipbon Pol.Eff.Dure 2011-01-01	cevious example d', ma.strings Acc.Date Par 2011-04-30 2011-07-01	is , uniclesses-, id. Indemnity Fer 0.00	Jalalypes, beader id. Expense Case.i 148.04	RSV. Indemnity Care. Re	207.4
fr fr ful ful ful ful ful ful ful ful ful ful	datatypes mpy data wad in ds sing/-res hesh keep sing blicy.ID poscoor poscoor poscoor	' already to clippo to d.teble(f rved deve Claim.ID : COSCOSI CHAP	defined in ; and, then ile="clipbon: Pol.Eff.Date 2011-01-02 2011-01-02	acc.Dete Par 2011-04-30 2011-07-01 2011-07-02	id.Indemnity Fat	detetypes, besder Ed. Expense Case. 148.04 57.14	RSV. Indemnity Care. Re NO. 60.52	207.4 90.5 71.0
o fin	datatype: mpy data wad in da sins/vice howk Assv sins blicy. ID poscool poscool poscool poscool	" already to clipho ta d.teble(f eved date Claim.ID C000001 C000003	defined in ; and, then ile="clipbon Pol.Eff.Date J011-01-01 F011-01-02 J011-01-03	Acc.Date Fa: 2011-04-90 2011-07-02 2011-07-02 2011-07-02	le, culClesser, culcless	detetypes, besder id.Expense Case.1 161.00 50.10 50.22 50. 25.02	85V. Indemnity Care. St 80. 60.52 64.69 80. 27.00	207.4 70.5 75.0
o fin	datatype: copy data wad in da sins/rite beat inser- sins blicy.ID poscoos poscoos poscoos poscoos	" already to clippo to d. Leble (f. greed dese Claim. ID : C000001 C000001 C000001	defined in ; hrd, then lle="clipbom: Pol_Eff_Dute 2011-01-01 2011-01-02 2011-01-04 2011-01-04	Acc.Date Fa: 2011-04-90 2011-07-02 2011-07-02 2011-07-02	id. Indemnity Fai 0.00 51.46 50.46	dstatypes,header id.Expense Case.i 161.04 57.16 50.22 90.	EdV. Indemnity Care. Re 10h 60.52 64.69 10h	207.4 70.5 71.0 9
o for our five columns of the column	datatype: opy data wad in da sing/-re- druk ingr aing olicy.ID poscool poscool poscool poscool poscool poscool	" already to dispose the dispose distance (Dispose cossosi cossosi cossosi cossosi cossosi cossosi	defined in ; and, then lite="clipton: pol.Eff.Date	Acc.Date Pai 2011-04-30 2011-07-02 2011-07-02 2011-07-02 2011-07-04 2011-07-04	id Indemnity Far 0.00 51.65 52.46 90.22.52 47.11	detetypes, besider id.Expense Case.1 168.00 50.10 50.10 90 25.00 60.00	85V. Indemnity Care. Re 80A 40.502 44.60 90A 27.60 74.40 37.00	207.4 70.5 71.0 8 20.8 85.2
0 #c 0 #z 0 ul 0 cl 0 cl 2 2 4 4 5	datatypes mpy data wad in da sinm'ura hewh keev sinm bolicy. ID poscool suusour Poscool cab- poscool suusour Poscool suusour Poscool	" already to clipto to di.teble(E eved deve Claim.ID C000001 c000004 C000004 C000005 C000006 C000006 C000006 C000006	defined in ; and, then ile="clipbon policy of the policy o	Acc.Date Fa: 2011-04-90 2011-04-90 2011-07-01 2011-07-02 2011-07-04 2011-07-08 2011-07-08	10	Jatatypes, header 16. Expense Case. 161.06 50.12 50. 25.00 60.01 16.00 15.46	85V. Indemnity Care. Re 85, 60.52 64.69 85, 27.60 76.68 17.02 61.65	207.4 70.5 71.0 80.8 86.7 19.7 90.7
) (1) (1) (1) (1	datatype: opy data wad in da sing/-re- druk ingr aing olicy.ID poscool poscool poscool poscool poscool poscool	" already to clipto to di.teble(E eved deve Claim.ID C000001 c000004 C000004 C000005 C000006 C000006 C000006 C000006	defined in ; and, then lite="clipton: pol.Eff.Date	Acc.Date Fa: 2011-04-90 2011-04-90 2011-07-01 2011-07-02 2011-07-04 2011-07-08 2011-07-08	id Indemnity Far 0.00 51.65 52.46 90.22.52 47.11	detetypes, besider id.Expense Case.1 168.00 50.10 50.10 90 25.00 60.00	85V. Indemnity Care. Re 80A 40.502 44.60 90A 27.60 74.40 37.00	207 70 71 20 85

Appendix

Packages

Several packages exist as alternate ways to exchange data with Excel (and Access)

- XLConnect: advanced manipulation of workbooks from within R, with key methods of readWorksheet, readNamedRegion, writeWorksheet, and xlcDump. Developed by Mirai Solutions.
 - http://cran.r-project.org/web/packages/XLConnect/vignettes/XLConnect.pdf
- xlsx: similar to XLConnect. Developed by Adrian A. Dragulescu.
 - http://cran.r-project.org/web/packages/xlsx/xlsx.pdf
- RODBC: allows connections via ODBC to MS Access, with key method sqlQuery. Developed by Michael Lapsley and Brian Ripley.
 - http://cran.r-project.org/web/packages/RODBC/RODBC.pdf

Exporting data from R

- $\bullet \quad \text{Basic command to copy data to clipboard: > write.table (claims, file= "clipboard")}\\$
- A couple extra specifications improve the experience
 ✓ Use option sep="\t" and the data will automatically be placed in separate columns (rather than having to perform a text-to-columns maneuver)

 ✓ Use option row.names=FALSE to remove numbering of the rows

 > write.table (claims, file="clipboard", sep="\t", row.names=FALSE)

- Basic command to write data to file: > write.csv (claims, file= "claims.csv")
- A couple extra specifications improve the experience
 Same options as write.table
- > write.csv (claims, file="claims.csv", sep="\t", row.names=FALSE)

_			
-			
-			
_			
-			
-			
_			
_			
_			
-			
-			