



Periodic Dataset 4

Importing Enroll-HD PDS Files

Enroll-HD

A worldwide observational study for Huntington's disease families

Enroll-HD PDS Reference Guides

A CHDI Foundation Project

Contents

1.	PURPOSE OF DOCUMENT	3
2.	DATA FILES PROVIDED ENROLL-HD	3
3.	IMPORT .csv FILES INTO EXCEL	3
4.	IMPORT .csv INTO R	12
5.	IMPORT Rdata FILES INTO R	12

1. PURPOSE OF DOCUMENT

This document provides instructions on how to import the .csv (comma-separated values) formatted Enroll-HD PDS data into Excel and R (a software environment for statistical computing and graphics). The document contains step by step instructions on how to open and format the data files to make the data easy to examine and use. There are other methods that may work equally as well. This document is not intended to be an exhaustive, but to simply provide one method.

2. DATA FILES PROVIDED ENROLL-HD

The Enroll-HD PDS dataset is provided in two formats:

- CSV file: CSV stands for commaseparated values (.csv) which is a delimiter-separated format. There are many types of delimiters including commas, semicolons, tabs, etc. The PDS data uses the tab as the delimiter (→ tab).
- R file: binary code format for the R software application.

The .csv file format can be imported into Excel spreadsheets as well as into most statistical software packages including R, Stata, and SAS.

In many cases it is essential to specify, before importing the data into the analysis software, that the variables are separated by tabs, as the default delimiter may change based on the country or region where the software was developed. It is also important that these files should not be edited in a word processing software or other programs that may potentially modify the tab characters, as this may damage the integrity of the original files. CSV files can be saved in other formats which are compatible with other statistical software packages as needed.

3. IMPORT .csv FILES INTO EXCEL

The .csv files can be easily imported and opened in Microsoft Excel. Because Excel is language dependent and delimiters differ from one country to another, some considerations need to be addressed when opening the .csv files to maintain data integrity. The procedures outlined here, to open the .csv files, can be applied to different versions of Excel.

As a default, Excel reads the values for each column as being in a "General" format. For example, unless otherwise specified, Excel interprets figures as numbers (e.g 1234), entered dates as date format (as pre-set, e.g. 11/28/2016), and it changes other values to text format (e.g. Aspirin). For some entries this is counterproductive, as Excel may misinterpret the entries and reformat the data, effectively changing the data (e.g. 1.5 is read as May 1 instead of 1.5 mg; or the WHO-DD Code for Tetrabenazine 00222101003 is changed to 22211003, removing the important leading "0"s). To maintain the integrity of the data, each data column needs to be carefully examined prior to importing the data into Excel.

Below are step by step guidelines for correctly importing the data into Excel:



Step 1 – Importing data: Either open .csv file using Excel, or open Excel then go to tab 'Data' and click on get external data 'From Text'. The file used for this demonstration is the 'profile.csv' file. In the first step, when the data is imported all data will be entered into the first column of the file.

	profile	- Microsof	ft Excel										_ 0	×
File Home Insert Page Layout Formulas Data Review View													∝ 🕜 🗆	er XX
Image: Second state of the se		i	?		· · ·	Show Detail Hide Detail	Data Analy	is						
From From From Other Existing Refresh	ove Data C	Consolidate Wh	nat-If Grou	up Ungroup S	Subtotal									
Access web Text Sources Connections All Connections	Data Tools	Anar	iysis * *	* Ou	tline	R	Δnalvsis							
A1 • • • • • • • • • • • • • • • • • • •	mbdmomagou	vdadbddadag	tocyfbycor	atreemtrage	evenhieve	mhddiagn	voctovratore	octofdovou	himoyfor	movrator				-
subjulegionsex acentation subsage displaced send agright agrowing mccdepccdepageccirbccirbageccvabccvabageccaptccaptageccpobccp	obageccpsyccp	syageccpsyfh	icccogccco	ogage	sksubjskie	anniuulagn	SACSISAIDICIS	esterusxst	Djiliskiali	inistiatei				-
A B C D E F G H I J	K	L	М	N	0	Р	Q	R	S	Т	U	V	W	
$\begin{tabular}{l} 1 \\ subjidregin{tabular}{l} subji$	cmtrccmtrages	sxsubjsxfamh	ddiagnsxes	stsxratersxe	stcfdsxsu	bjmsxfamn	nsxratermccd	epccdepag	eccirbccir	bageccvab	ccvabageco	aptccaptag	eccpobccp	opt
2 R000024542Northern Americam110441915001142424247142111101460014600														
3 R00021725XEuropef1103824155010NANANANA0000000														
4 R00023946XEuropem110412001701165677070166211101NA1710000														
5 R000798865Europef1104420013010NANANA0000000														
6 R001084542Northern Americaf1122151														
7 R001159219Europef1101716014010000000														
8 R001225384Europem1104117013801454544461441226145144001440144														
9 R001286120Northern Americaf6104617132000NANANANANANA1701150000														
10 R00170741XEuropem1104718125011354035401351111013801411400141														
11 R001717535Northern Americam11127171560100000012300														
12 R001832046Europem111421813500165656568165111116800017000														
13 R001924217Northern Americam8123160														
14 R002201087Northern Americaf1204414140011494949521492111153153014515300														
15 R002215749Northern Americaf1101817150000NA000013100														
16 R002311975Northern Americaf1510401600116116116100000														
17 R002330616Europem1104116150010000000														
18 R00257048XEuropem1120161														
19 R002593663Europem1104220160010NANANANANA0000000														
20 R002776052Europef1123171														
21 R002798312Northern Americaf11042210140115454NANA1NANA0000000														
22 R002834769Europem110432001501151484852148233314800000154														
23 R002901876Northern Americaf11020190143100000000														
24 R002918968Northern Americaf2111616014410NANA0137000000														
25 R002995975Northern Americam1118181														
26 R002999525Northern Americaf11145171400101251231230000														
27 R00300506XEuropef11142161420114444NA45145166613914814913814800														
28 R003027006Northern Americaf1104326149010NANANANANANA000000135														
29 R003114308Europef110451601531135353538135111114515315314101530138														
30 R003161464Europef110451901401131311312NANA101100110000														
31 R003183902Northern Americaf120371815501154545401210000000					_									
prome / W				1					III			1000		
reauy												100%		J



Step 2 – Separating data: Select the first column and on the tab 'Data' click on 'text to column'. A block screen – "wizard" will appear. This wizard will help guide you to proper separation of the data.

File Home Insert Page Layout Formulas Data Review View Image: Construction of the participant of the partipant of the partipant
Image: Construction of the construc
Cat External Data Connections Sort 8: Eller Data Tools Outline E Analysis
At the subjidregionseyracebandedhysiddssagedsplacedsendcadh Tard to Columna biddadagesyfhycomtrocmtragesysubjisyfamhddiagnsyestsyratersyestofdsysubjimsyfammsyrater
AL
A B C D F F G H Cellino separate columnis I M N O P O R S T U V W C
subjidregionsexracehandedhssiddssagedsplacedsendcaghighcaglowmomhdmomagesxdadhd Roperate a Ropolog 2522Northern Americam110441915001142424247142111101460014600 column of full names into separate rol separate a rol manufactorial separate a rol
3 R0002172\$XEuropef1103824155010NANANANAN0000000 first and last name columns.
4 R0002394\$XEuropem110412001701165677070166211101NA1710000 In Word, use this feature to convert
5 R0007988%5Europef1104420013010NANANA0000000 splitting the twin to a table,
6 R001084542Northern Americaf1122151 each comma period, or other
7 R0011592i 9Europef11017160140100000000 character you specify.
8 R001225334Europem1104117013801454544461441226145144001440144 Image: Press F1 for more help.
9 R001286120Northern Americaf6104617132000NANANANANA1701150000
10 R00170741XEuropem1104718125011354035401351111013801411400141
11 R00171753 SNorthern Americam1112717156010000012300
12 R0018320456Uropem111421813500165656568165111116800017000
13 R0019242L /Northern Americant81221160
14 R00220108/Northern America120441414001149494952149211153153014515300
15 R0022157491Northern America11018171500010A000013100
10 RV023119 SNorthern America 11510401001110110110100000
17 NOV25300 Deuropeni104105300 100000000
10 N002502632Europem10/0220150010NANANANANANANANANANANANANANANANANANA
21 R002798312Northern Americaf11042210140115454NANA1NANA0000000
22 R002834799Europem110432001501151484852148233314800000154
23 R0029018 6Northern Americaf1102019014310000000
24 R0029189§8Northern Americaf2111616014410NANA0137000000
25 R0029959 SNorthern Americam1118181
26 R0029995 SNorthern Americaf11145171400101251231230000
27 R0030050\$XEuropef11142161420114444NA451451656613914814913814800
28 R00302700 Konthern Americaf1104326149010NANANANANANANA000000135
29 R003114308Europef110451601531135353538135111114515315314101530138
30 R003161494Europef110451901401131311312NANA101100110000
31 R003183902Northern Americaf120371815501154545401210000000
Ready Count: 8847 Image: Count: 8847 Image: Count: 8847 Image: Count: 8847



Step 3 – Choose way of separating the data: In the 'text to column' wizard, select the 'delimited' option; this will separate the variables in separate columns and then click 'Next'.

X	10 -	(°" ~ -													prot	file - Mici	osoft Ex	cel										_ 🗇 🔿	ĸ
Fil	- F	lome	Insert	Page Lay	out F	ormulas	Data	Review	View																			a 🕜 🗆 🖻	J 23
Â	1		P			62	D Conne	ctions	↓ A		W Cl	lear eapply	*****		Ø		2	¢[t			● Show Detail ■ Hide Detail	Data Analysi	is						
Fron	From	From	From Oth	ner Ex	isting	Refresh	🖙 Edit Li	nks Z	↓ Sor	t F	ilter 🗸 A	dvanced	Text to	Remove	Data	Consolidat	e What-If	Gro	up Ungrou	p Subtotal									
Acces	s web	Get Ext	ernal Data	· Con	nections	All *	onnections			Sort	& Filter	arancea	Columns	Duplicates	Data To	ols	Analysis *			Outline	5	Analysis							
	Δ1		• (n		fr sub	iidregia	onsexrace	ehanded	hxsidd	ssager	Isplaceds	endcagł	ighcagle	owmomh	dmoma	resxdadhdo	ladagesxf	hxccn	otreemtra	resysue	isxfamhddiag	isvestsvratersv	estofdsxs	uhimsxfa	mmsyrater				
	712				mc	cdepcc	depageco	irbccirb	ageccy	abccv	abagecca	ptccapt	ageccpc	bccpoba	geccosv		psyfhcccc	gccco	ogage	Beened	joxramiaalaB.	is to start at crow		abjinoxia					
	А		В	С	D		F	F	<u> </u>	G	н	· ·	1		к	170	N	1	N	0	р	0	R	S	Т	U	V	W	
1	subiidre	gionsex	racehan	dedhxsi	iddssage	edsplac	- edsendca	Ighighca	glowm	omhd	momages	xdadhd	dadages	xfhxccm	trccmtra	gesxsubisx	amhddia	gnsxe:	stsxraters	sxestcfds	sxsubimsxfam	msxratermccde	epccdepag	eccirbcc	irbageccva	bccvabagec	captccaptag	eccpobccpol	bi
2	R00002	4542Nc	rthern A	merica	m11044	19150	01142424	247142	11110:	4600	14600					0,					, ,			,	Ŭ				
3	R00021	725XEu	ropef11(038241	55010N	ANANA	NANAOO	00000																					
4	R00023	946XEu	ropem11	041200	01701			Com	ort T	out to	Column		rd Ct	op 1 of 1	,	2 X													
5	R00079	8865Eu	ropef11(044200	13010			CON	vent n		Column	IS VVIZA	ru - su	epitora	>														
6	R00108	4542Nc	rthern A	merica	f1122	The Text	Wizard has	determin	ed that y	our dat	a is Delimite	d.																	
7	R00115	921.9Eu	ropef11(017160	14010	If this is	correct, cho	ose Next, (or choos	e the da	ata type that	best des	ribes your	r data.															
8	R00122	5384Eu	ropem11	041170	01380	Origina	l data type																						
9	R00128	6120Nc	rthern A	merica	f6104	Choose	e the file typ	e that bes	t describ	es your	data:																		
10	R00170	74 XEu	ropem11	047181	12501	۲	Delimited	- Chara	octers su	ch as co	ommas or ta	bs separa	te each fie	eld.															
11	R00171	7535Nc	rthern A	merica	m111	0	Fixed width	n - Fields	are alig	ned in o	olumns with	n spaces b	etween ea	ach field.															
12	R00183	2046Eu	ropem11	142181	13500																								
13	R00192	4217Nc	rthern A	merica	m812																								
14	R00220	1087Nc	rthern A	merica	f1204																								
15	R00221	5749Nc	rthern A	merica	f1101	Preview	w of selecte	d data:																					
16	R00231	1975No	rthern A	merica	f1510																								
17	R00233	0616Eu	ropem11	041161	15001	1 sub 2 R00	0024542N	nsexrace orthern	Ameri	ihxsid cam110	441915001	placeds 1142424	endcaghi 2471421	ighcaglov	mom														
18	R00257	048XEu	ropem11	20161		3 R00	021725XE	uropef11	103824	155010	NANANANAN	NAOOO																	
19	R00259	3663Eu	ropem11	042201	16001	5 R00	023946XE 0798865E	uropemi: uropefii	L04120	017011 013010	NANANA000	1662111 DO)																
20	R00277	6052Eu	ropef112	23171		6 R00	1084542N	orthern	Ameri	caf112	2151					Y													
21	R00279	8312Nc	rthern A	merica	f1104	`																							
22	R00283	4769Eu	ropem11	.043200	01501					Ca	ncel	< Ba	:k	Next >		Finish													
23	KUU290	1876Nc	rthern A	merica	1102	_		_																					
24	R00291	8968NC	rthern A	merica	211101	00144	TOWAWAG	137000	000								_												
25	KUU299	S975NC	rthern A	merica	m11181	81	10125122	112000	0																				
20	100299	SOSTE	ronof11	merica	12011451	1400		6120140	0	01/00	0																		
27	00200		oper11.	L421014	+201144	61400	100100100	0139140	NA000	000113	5																		
28	200211	12005	ronern A	merica	1521125	01490			1E214	10152	0120																		
29	200214	4308EU 1464E…	roper110	745100.	1/01124	21121	20132111	110011	13314	10123	0139																		
21	000010	20020	oper110	743130.	1401131 F120274	01550		01210011	0000																				-
	▶ N F	profile		merica	1705/1	61.3.30	1.040404	012100											I] ◀						· · · · · · · · · · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·		
Read	/																								Count: 8847	ШПШ	100% (-)		(+)



Step 4 – Choose delimiter: Selecting the "Delimited" radio button lets Excel know that the data is separated by commas or tabs. Once you click next the wizard will provide a preview of how the data will be separated into different columns..Then click "Next" to proceed.

🛣 🚼 🔊 - 🝽 - 🖛 _ 🗇 🗡 profile - Microsoft Excel Data Review a 🕜 🗖 🗗 🛛 Home Insert Page Lavout Formulas View 📲 Show Detail 🛛 💾 Data Analysis Connections * 🐝 Clear 22 **≙**↓ A Z A 변불 A -■ Hide Detail Properties Reapply X↓ Sort From From From Other Existing Refresh Filter Text to Remove Data Consolidate What-If Group Ungroup Subtotal 📟 Edit Links Advanced Columns Duplicates Validation * Access Web Text Sources Connections All -Analysis * Get External Data Sort & Filte Data Tools Outline Analysis **v** (e) fx subjidregionsexracehandedhxsiddssagedsplacedsendcaghighcaglowmomhdmomagesxdadhddadagesxfhxccmtrccmtragesxsubjsxfamhddiagnsxestsxratersxestcfdsxsubjmsxfammsxrater A1 mccdepccdepageccirbccirbageccvabccvabageccaptccaptageccpobccpobageccpsycpsyageccpsyfhcccogcccogage А B C D F F G H I J Κ М 0 Ρ 0 R U V W nsexracehandedhxsiddssagedsplacedsendcaghighcaglowmomhdmomagesxdadhddadagesxfhxccmtrccmtragesxsubjsxfamhddiagnsxestsxratersxestcfdsxsubjmsxfammsxratermccdepccdepageccirbccirbageccvabccvabageccaptccaptageccpobccpob subjidregio R000024542Northern Americam110441915001142424247142111101460014600 R00021725XEuropef1103824155010NANANANANA0000000 R00023946XEuropem110412001701 ? × Convert Text to Columns Wizard - Step 2 of 3 R000798865Europef1104420013010 R001084542Northern Americaf1122 This screen lets you set the delimiters your data contains. You can see how your text is affected in the preview below. R001159219Europef1101716014010 R001225384Europem110411701380 Delimiters ✓ <u>T</u>ab 9 R001286120Northern Americaf6104 Treat consecutive delimiters as one 10 R00170741XEuropem110471812501 Semicolon Comma R001717535Northern Americam111 11 Text gualifier: " \sim R001832046Europem111421813500 Space 12 Other: 13 R001924217Northern Americam812 R002201087Northern Americaf1204 14 Data preview 15 R002215749Northern Americaf1101 R002311975Northern Americaf1510 16 dsend cagh ^ subiid region anded hxsid dssage dsplace R002330616Europem110411615001 sex race 17 R000024542 Northern America 44 38 R00257048XEuropem1120161 18 R00021725X Europe R00023946X Europe 41 44 19 R002593663Europem110422016001 00798865 Europe v 20 R002776052Europef1123171 R001084542 Northern America 22 21 R002798312Northern Americaf1104 22 R002834769Europem110432001501 <u>F</u>inish Cancel < Back <u>N</u>ext > 23 R002901876Northern Americaf1102 R002918968Northern Americaf2111010014410INAINA015700000 24 25 R002995975Northern Americam1118181 R002999525Northern Americaf11145171400101251231230000 26 27 R00300506XEuropef11142161420114444NA45145166613914814913814800 28 R003027006Northern Americaf1104326149010NANANANANANA000000135 29 R003114308Europef110451601531135353538135111114515315314101530138 R003161464Europef110451901401131311312NANA101100110000 30 31 R003183902Northern Americaf120371815501154545401210000000 🛚 🔹 🕨 🚽 profile ⁄ 😓 4 Ready Count: 8847 🔠 🗐 🗐 100% 😑



Step 5 – Assign formats: For each column a format needs to be assigned. The default format 'General' works for most columns. Columns where numbers have leading "0" and columns with mixed entries like 1.5, 1,5, 1/5, need to be explicitly formatted as Text, as entries might otherwise become corrupted in an unchangeable way. After assigning the format to each column click 'Finish'.

X	profile - Micro	osoft Excel					_ 🗇 ×
File Home Insert Page Layout Formulas Data Review View							X 🖬 🗕 🜔 A
From Access From Web From Other Text Existing Connections Existing Connections Properties All - Properties - XI XI	K Clear Clear Image: Clear	What-If Analysis Couting		Data Analysis			
A1 - f subjidregionsexracehandedhxsiddss	agedsplacedsendcaghighcaglowmomhdmomagesxdadhdda	adagesxfhxccmtrccmtrages	xsubjsxfamhddiagn	sxestsxratersxestcfds	xsubjmsxfammsxrater		~
mccdepccdepageccirbccirbageccva	occvabageccaptccaptageccpobccpobageccpsyccpsyageccp	syfhcccogcccogage					-
A B C D E F	S H I J K L	M N	O P	Q R	S T	U V	w 📮
1 subjidregionsexracehandedhxsiddssagedsplacedsendcaghighcaglowmo	mhdmomagesxdadhddadagesxfhxccmtrccmtragesxsubjsxfa	amhddiagnsxestsxratersxes	tcfdsxsubjmsxfamm	sxratermccdepccdep	ageccirbccirbageccvabc	cvabageccaptccap	tageccpobccpob
2 R000024542Northern Americam1104419150011424242471421111014	60014600						
3 R00021725XEuropef1103824155010NANANANANA0000000							
4 R00023946XEuropem110412001701 Convert Tex	t to Columns Wizard - Step 3 of 3 ? 🛛 🕴						
5 R000798865Europef1104420013010							
6 R0010845 2Northern America f1122 This screen lets you select each column a	nd set the Data Format.						
7 R001159219Europef1101716014010 Column data format							
8 R001225334Europem110411701380	eneral' converts numeric values to numbers, date values to dates, and all						
9 R001286120Northern Americaf6104rextre	maining values to text.						
10 R00170741XEuropem110471812501 O Date: DMY V	Advanced						
11 R001717535Northern Americam111 Do not import column (skip)							
12 R001832046Europem111421813500							
13 R00192421/Northern Americam812 Destination: \$A\$1							
14 R00220103/Northern America1204 Data preview							
15 R002213/19Northern America11101							
17 P002220616Euronem110/11615001 Subjid region	GenerGenerGeneralGeneraGeneralGeneral GeneraGene sex race handed hxsid dssage dsplace dsend cagh ^						
17 R002550010Edropem1120151 18 R000024542 Northern America	m 1 1 0 44						
19 R002593663Europem110422016001 R00021725X Europe	m 1 1 0 38						
20 R0027760%2Europef1123171							
21 R002798312Northern America1104	>						
22 R002834769Europem110432001501							
23 R0029018 6Northern Americaf1102	Cancel < Back Next > Finish						
24 R002918958Northern Americaf211101001441018AIVA0157000000							
25 R002995975Northern Americam1118181							
26 R002999525Northern Americaf11145171400101251231230000							
27 R00300506XEuropef11142161420114444NA451451666139148149138	14800						
28 R003027006Northern Americaf1104326149010NANANANANANA0000	00135						
29 R003114308Europef11045160153113535353813511111451531531410	1530138						
30 R003161464Europef110451901401131311312NANA101100110000							
31 R003183902Northern Americaf120371815501154545401210000000							▼
Peady Pady					Count: 00.47		
Neauy					Count. 8847		

$\left(\right)$	Version	1)

NOTE: The data files pharmacotx and nutsuppl contain two columns 'cmtrt_decod' and 'cmdostot' that require formatting as Text.

X	Image: Image of the second												icrosoft E	xcel									_ □	× ar x
Fr	om From cess Web	From Fro Text So Get External	m Other E urces ~ Co Data	Existing nections A	fresh Connections	tions ties ks X↓	Sort F	W Clea k Clea k Rea Adva & Filter	r pply Tex anced Colu	t to Remove	Data s Validation Data Too	Consolidate	What-If Analysis *	Group Ungr	oup Subtotal	●∃ Show Detail ■∃ Hide Detail	Data An	alysis s						
	A1	•	(*	<i>f</i> ∞ subjid	cmtrtmodif	ycmtrtc	ecodcmtr	t_ingcmtrt	atccmir	ndcmodif	ycmindc	decodcmd	ostotcmd	ose_cmdo	sucmdosfro	cmroutecms	tdycmenrfo	cmendy						*
	A	В	С	D	E	F	G	Н	1	J	К	L	M	N	0	Р	Q	R	S	Т	U	V	W	=
1	subjidcm	trtmodi	fycmtrtd	lecodcmtrt_	_ingcmtrta	tccmindc_		mindcdec	odcmdos	totcmdose_	_cmdosuc	mdosfrqcm	nroutecms	tdycmenrfo	mendy									
2	R000024	542Xenazi	ne0022210	1003Tetrab	enazineN07X)	(ChoreaD	10008748	37.5mg110	0283															
3	R000024	542Aleve0	025620201	18Naproxen	sodiumG02C	C,M01AE,I	M02AABa	ck painD.10	00398844(0mg11-4001														
4	R000024	542Xenazi	ne0022210	1003T		Convo	t Toyt to	Columns	Wizard	Stop 2 of	2	? ×												
5	R000024	542Celexa	005826020	006Cita		COnve	IT TEXT TO	Columns	wizaru -	step 5 01	5													
6	R000024	542Clonaz	epam0028	52010 This	s screen lets you s	select each c	olumn and s	et the Data For	mat.															
7	R000798	8650ki003	21702005	Ketopr Co	lumn data format	-																		
8	R000798	865Tachip	irina00020	00104) <u>G</u> eneral		Conor	al' convorte pur	ooric valuos i	to numbero, da	to values to s	atoc and all												
9	R001084	542Citalop	oram00582	60100	Text		remain	ing values to te	xt.	to numbers, ud	te values to t	ates, and an												
10	R001084	542Simvas	tatin00848	310100	<u>D</u> ate: DMY	~	1		Ad	vanced														
11	R001084	542Symbic	ort015381	.01002) Do not import c	olumn (skip)	-																	
12	R001159	21.9Simvas	tatin00848	310100																				
13	R001225	384Rhinoc	ort hayfev	er006: D <u>e</u> s	tination: \$A\$1							1												
14	R001225	384Parace	tamol0002	200010 Da	ta <u>p</u> review																			
15	R001225	384Ibupro	fen001092	05015																				
16	R001225	384Escital	opram0158	88501(<u>G</u>	eneral Ger	eral	Text	Gene	ral		General	_												
17	R001707	41XSulpirio	00314301	039Sul R	ubjia cmt 000024542 Xer	azine	0022210	1003 Tetr	ting abenazine		Cmtrtat													
18	R001707	41XTiaprid	004357020	D36Tia R	000024542 Ale	ve	0025620	2018 Napr	oxen sodi	um	G02CC,M01	AE, MO												
19	R001707	41XSertral	in0101140	1011S	000024542 Ker 000024542 Cel	azine lexa	0022210	02006 Cita	abenazine lopram hy	drobromide	NU/XX NO6AB													
20	R001707	41XSertral	in0101140	1011S R	000024542 Clo	nazepam	0.028520	01001 Clon	azepam		NO3AE.NO5	BA												
21	R001707	41 XInvesti	gational dr	ug999	•							,	-29											
22	R001707	41XLamot	rigine0104	71010			Ca	ncel	< Back	Next 2	>	Finish												_
23	R001707	41.XMelpe	rone00219	/0100					-			-												
24	R001707	4 Xinvesti	gational dr	ugoooaaa	1001IIIvestiga	cionai uruį	, ingreuie	nt unspection		pathyp.100	02942500	IIG11-4200	-305											
25	R001707	4 XInvesti	gational dr	ugaaaaa/01	LUUIINVestiga	tional drug	, ingredie	nt unspecifi	edv03AXA	pathyD.100	02942150	ng11-3640	-352											
20	R001707	4 Ainvesti	gational dr	n8aaaaaa101	1001 Investiga	tional drug	, ingredie	nt unspecifi	advozaXA	pathyD,100	029421501	ng11-3510	-337											— I
2/	R001707	4 Anvesti	gational dr	neoooooooo	1001 Investigation	tional drug	, ingredie	nt unspecifi	advozava	pathyD 100	029423001	ng11-3300	-280											
28	P001707	4LAINVest	Barrouga di	ugaaaaaa/01	LOOTIIIVestiga	Sloop disc	, ingreuie rdorD 100	nt unspecifi	-uvu3AXA -112651	ратнур. 100	029421301	IIBTT-5130	-215											
29	R001822	046Tetrab	anazine00219	2221010027	etrahenazino		rucio.100	001301619	75mg11 7	760-550														
30	R001822	046Loraze	nam00272	2010011 019		NOSBAAR	vietyD 10	001391018.	og11_4701	100-330														-
M	< ► ► pl	harmacotx		2010011018	CCUALINUSAF		AIGLVD. IU		is 11-4771															
Re	ady																		C	ount: 39483	田 田 田	100% (-)		(+)



Step 6: The .csv file is column-separated and should be saved as an Excel file (.xls or .xlsx) using the 'Save As' option.

	🛛 🖬 🤊 • 🖓 • 🖓 •													×										
Fi	ile Ho	me Insert	Page Lay	out Formu	las Data	Review	View																۵ 🕜 ه	- # X
Fro	m From	From From	Other Ex	isting Refr	esh	erties	↓ <mark>Z Z</mark> ↓ Sort Fi	Itter	oply Tex	xt to Remove	Data (Consolidate	What-If C	iroup Ungrou	ip Subtotal	Show Detail Hide Detail	Data An	alysis						
ACCE	255 Web	Get External [ata	lections All	Connection	IS	Sort	& Filter	Con	umns Duplicates	Data Tools	A	naiysis *		Outline	5	Analysi	s						
	A1	-		f _æ subiid																				
																								*
	А	В	С	D	E	F	G	Н	1	J	K	L	М	N	0	Р	Q	R	S	Т	U	V	W	
1	subjid	region	sex	race	handed	hxsid	dssage	dsplace	dsend	caghigh	caglow	momhd	momage	s> dadhd	dadagesx	fhx	ccmtr	ccmtrage	sxsubj	sxfam	hddiagn	sxest	sxrater	SX€ ≡
2	R000024	5 Northern	/m	1	1	1	0			44	19	1	. 5	0 (0	1	l 1	42	2 4	42 42	2 47		1 /	42
3	R000217	2.Europe	f	1	1 1	1	0			38	24	1	. 5	5 (0	1	L C)	NA	NA				
4	R000239	4 Europe	m	1	1	1	0			41	20	0)		1 70) 1	1 1	65	5 (57 70) 70		1 (66
5	R000798	8 Europe	t	1		1	0			44	20	()		1 30) 1)	NA	NA		NA		
6	R001084	Northern	/f	1		1	0			22	15]								
/	R001159	2.Europe	1	1		1	0			1/	10) \		1 40		L L	, 46	:	16 4/	1 46		1	44
0	R001223	1 Northorn	111 / F	1		1	0			41	17	1)	2 (1 30			43		+5 44 NA	+ 40	NIA	1 '	44
10	P001200	1 Europe	m	1		1	0			40	10	1	. 3	5 (0	1) (1	20	11/24	10 25	5 40	INA	1	25
11	R001717	5 Northern	/m	1		1	1			27	10	1	. 2	6 1	0	1		1 30	, .	+0 5.	, 40		<u> </u>	35
12	R001832	0 Europe	m	1		1	1			42	18	1	3	5 (0) 1	, 6'	5	55 65	5 68		1	65
13	R001924	2 Northern	4m	2	3 1	1	1			23	16			5 .	0	()		, ,		,			
14	R002201	0 Northern	/ f	1		2	0			44	14	1	4	0 0	0	1	I 1	49)	49 49	9 52		1	49
15	R002215	7 Northern	<i>k</i> f	1		1	0			18	17	1	5	0 (0	() ()	NA				-	
16	R002311	9 Northern	∕f	15	5	1	0			40	16	C)	(0	1	L 1	61	L					
17	R002330	6 Europe	m	1	1 1	1	0			41	16	1	. 5	0 (0	1	L C)						
18	R002570	4 Europe	m	1	1 1	1				20	16					1	L							
19	R002593	6 Europe	m	1	1 1	1	0			42	20	1	6	0 (0	1	ι α)	NA	NA				
20	R002776	0.Europe	f	1	L 1	1				23	17					1	L							
21	R002798	3 Northern	∕f	1	1 1	1	0			42	21	C)		1 40) 1	L 1	54	l !	54 NA		NA		
22	R002834	7 Europe	m	1	L 1	1	0			43	20	C)	:	1 50) 1	l 1	51	L 4	48 48	3 52		1	48
23	R002901	8 Northern	/f	1	L 1	1	0			20	19	C)	:	1 43	: 1	ιc)						
24	R002918	9 Northern	/f	2	2 1	1	1			16	16	C)	:	1 44	4 1	L C)	NA	NA			0	
25	R002995	9 Northern	/m	1	1 1	1				18	18					1	L							
26	R002999	5 Northern	4 f	1	1 1	1	1			45	17	1	. 4	0 (0	1	L C)						
27	R003005	OEurope	t	1	L 1	1	1			42	16	1	. 4	2 (0	1	L 1	L 44	4	14 NA	45		1 4	45
28	R003027	0 Northern	/t	1	L 1	1	0			43	26	1	. 4	9 (0	1	L C)	NA	NA		NA		
29	R003114	3 Europe	t c	1	L 1	1	0			45	16	0)		1 53	1	L 1	1 35		35 35	38		1	35
30	R003161	4 Europe	T	1		1	0			45	19	0	-	-	1 40) 1	L 1	. 31			31		1	31
31	IR003183	ofile	μT	1			0			37	18	1	5	а (4	1	1	54		54 54	+		U .	
Read	dy																			Count: 8847	₩□ Ш	100% 🗩		+

4. IMPORT .csv INTO R

First, a package capable of reading CSV files has to be loaded into R environment. For example "readr" is one of the most popular applications. If a package like "readr" is not already installed, the CVS data files can be imported using the following code line: **install.packages (readr)**. Then, to load the CSV data into R using a package like "readr" use the: **library (readr)** command. To ensure the CSV file is imported correctly, set the directory to the file folder where the PDS files are located, and then run the following code:

file = read_delim("file.csv", "\t", escape_double = FALSE, trim_ws = TRUE)

5. IMPORT R data FILES INTO R

This data file is specific for R. After loading the R data files into R, 9 data frames are made available in the R environment and are ready to be used. The loading can be done using the function **load("Rdata_directory")**. For Rstudio users, the loading can be performed by clicking in the "load workspace" ribbon, and then browsing for the location of the R data file.



Importing Enroll-HD PDS files | PDS4

Revision History

Document Name	Summary of Changes
Version 2015-10-R1	Initial version for second Enroll-HD periodic dataset
Version 2016-10-R1	Revised version for third Enroll-HD periodic dataset
Version 2018-10-R1	Third version for fourth Enroll-HD periodic dataset

Importing Enroll-HD PDS files | PDS4