

Common Issues in Data Formats Submissions

CVB Statistics

March 2, 2016

Properly formatted data is suitable for immediate use in other software without need for additional manipulation that may delay analysis or introduce errors.

Tables and screenshots in this document addresses common issues that occur when formatting data tables for submission as one of the CVB data formats. More detailed information is given in the instructions for specific data formats at: https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/veterinary-biologics/biologics-regulations-and-guidance/ct_vb_data_formats_overview.

DO	DON'T
Start data entry in first row, first column of table (position 'A1' in Excel, OpenOffice)	DON'T Include empty rows and/or columns
Indicate missing data by NA	DON'T Indicate missing data by 'NA', na, n/a or N/A
Report raw dilution values	DON'T Report log dilution values DON'T Report dilution values as ratios '1:10', '1/10'
Include an entry (data or NA) for every cell in each row and column of the table	DON'T Leave cells empty DON'T Merge cells
Follow nomenclature outlined in "Creating .CSV Files for CVB Statistics Data Formats"	DON'T Include multiple tables in a single file
Report raw data once per observation	DON'T include raw data multiple times if different analysis is performed

Example

Start table in upper left position.

Use NA to indicate missing data values

ID	Tgroup	Treatment	Route	Litter	Pen	Sstatus	Day	Date	La	Lc	Ld	Ra	Rc	Rd	Int
278	A	vac	IM	4775	1	live	14	10/16/2010	0.3	0	0	1.1	0.4	0.0	0.0
309	A	vac	IM	4775	5	live	14	10/16/2010	2	1.5	7.1	1.2	1.0	0.6	0.0
297	A	vac	IM	4775	7	live	14	10/16/2010	0.1	0	0	0.6	1.2	0.0	0.0
288	A	vac	IM	5813	2	live	14	10/16/2010	0	2.4	1.4	5.2	0.9	0.0	0.0
311	A	vac	IM	5813	6	live	14	10/16/2010	0	0	0	0	0	0.0	0.0
301	A	vac	IM	7394	3	live	14	10/16/2010	2.6	3.1	0	0.3	0.3	0.0	0.0
306	A	vac	IM	4759	4	live	14	10/16/2010	0	0.3	0	1.4	0.0	0.0	0.0
277	A	vac	IM	4759	4	dead	11	10/13/2010	0	0	0	0.6	3.4	0.0	0.0
308	B	vac	SC	4775	1	live	14	10/16/2010	0	0.4	0.9	0	0	0.0	0.0
282	B	vac	SC	4775	5	dead	11	10/13/2010	0.5	0.8	3.3	0.4	0.6	0.0	0.0
295	B	vac	SC	4775	7	live	14	10/16/2010	0.2	0	0	0.9	0.0	0.0	0.0
220	B	vac	SC	5813	2	dead	12	10/14/2010	0	0	0	1	0.0	0.0	0.0
310	B	vac	SC	5813	6	live	14	10/16/2010	3.8	4.1	23.5	5.6	2.3	2.7	0.0
225	B	vac	SC	7394	3	live	14	10/16/2010	2.2	0.4	11.2	1.1	0.0	0.0	0.0
276	B	vac	SC	4759	4	live	14	10/16/2010	0	3.2	11.6	4.6	0.6	0.0	0.0
293	B	vac	SC	4759	4	dead	13	10/15/2010	0	0.3	0.9	0	0.0	0.0	0.0
298	C	pla	IM	4775	1	dead	12	10/14/2010	2.3	0	6.4	0.2	0.0	0.0	0.0
217	C	pla	IM	4775	5	dead	12	10/14/2010	4.7	5.1	31.3	1	0.5	0.0	0.0
307	C	pla	IM	4775	7	dead	11	10/13/2010	0	0	7.8	8.9	0.0	0.0	0.0
289	C	pla	IM	5813	2	live	14	10/16/2010	4.7	5.8	20.1	6	5.0	23.3	4.0
223	C	pla	IM	5813	6	dead	13	10/15/2010	0.8	1.4	1.7	2.5	2.1	1.2	1.6
314	C	pla	IM	7394	3	dead	12	10/14/2010	1	1	6.9	1.6	2.0	8.5	2.8
328	C	pla	IM	4759	4	dead	10	10/12/2010	2.8	3.2	5.9	4.8	0.5	9.3	2.4
216	D	pla	SC	4775	1	dead	12	10/14/2010	3.5	2.3	13.3	2.3	2.2	14.2	2.4
283	D	pla	SC	4775	5	live	14	10/16/2010	3	2	6.4	3.4	1.2	10.8	1.9
214	D	pla	SC	4775	7	dead	12	10/14/2010	3.2	2.7	21.4	7.7	9.5	23.3	3.2
300	D	pla	SC	5813	2	dead	10	10/12/2010	1.7	2.9	5.7	3.9	4.0	0.0	0.0
299	D	pla	SC	5813	6	dead	13	10/15/2010	2.8	5	8.9	5.1	2.6	0.0	0.0
290	D	pla	SC	7394	3	live	14	10/16/2010	2.8	2.9	11.6	4.8	4.3	0.0	0.0

(a) Right

WRONG

Empty rows and columns may cause data imports to be incomplete

Use of improper format for NA requires manual correction.

Including text in a numeric column requires further investigation and/or manual correction

Merged information will be lost when converting to CSV

ID	Tgroup	Treatment	Route	Litter	Pen	Sstatus	Day	Date	La	Lc	Ld	Ra	Rc	Rd	Int
278	A	vac	IM	4775	1	na	14	10/16/2010	0.3	0	0	1.1	0.4	0.0	0.0
309	A	vac	IM	4775	5	N/A	14	10/16/2010	2	1.5	7.1	1.2	1.0	0.6	0.0
297	A	vac	IM	4775	7	live	14	1	0.1	0	0	0.6	1.2	0.0	0.0
288	A	vac	IM	5813	2	live	14	10/16/2010	0	2.4	1.4	5.2	0.9	0.0	0.0
311	A	vac	IM	5813	6	live	14	10/16/2010	0	0	0	0	0	0.0	0.0
301	A	vac	IM	7394	3	live	14	10/16/2010	2.6	3.1	0	0.3	0.3	0.0	0.0
306	A	vac	IM	4759	4	live	14	10/16/2010	0	0.3	0	1.4	0.0	0.0	0.0
277	A	vac	IM	4759	4	dead	11	unknown	0	0	0	0.6	3.4	0.0	0.0
282	B	vac	SC	4775	1	live	14	10/16/2010	0	0.4	0.9	0	0	0.0	0.0
282	B	vac	SC	4775	5	dead	11	10/13/2010	0.5	0.8	3.3	0.4	0.6	0.0	0.0
295	B	vac	SC	4775	7	live	14	10/16/2010	0.2	0	0	0.9	0.0	0.0	0.0
220	B	vac	SC	5813	2	dead	12	10/14/2010	0	0	0	1	0.0	0.0	0.0
310	B	vac	SC	5813	6	live	14	10/16/2010	3.8	4.1	23.5	5.6	2.3	2.7	0.0
225	B	vac	SC	7394	3	live	14	10/16/2010	2.2	0.4	11.6	4.6	0.6	10.8	1.4
276	B	vac	SC	4759	4	live	14	10/16/2010	0	3.2	11.6	4.6	0.6	0.0	0.0
293	B	vac	SC	4759	4	dead	13	10/15/2010	0	0.3	0.9	0	0.0	0.0	0.0
298	C	pla	IM	4775	1	dead	12	10/14/2010	2.3	0	6.4	0.2	0.0	5.6	0.8
217	C	pla	IM	4775	5	dead	12	10/14/2010	4.7	5.1	31.3	1	0.5	21.8	4.0
307	C	pla	IM	4775	7	dead	11	10/13/2010	0	0	7.8	8.9	0.0	13.2	2.1
289	C	pla	IM	5813	2	live	14	10/16/2010	4.7	5.8	20.1	6	5.0	23.3	4.0
223	C	pla	IM	5813	6	dead	13	10/15/2010	0.8	1.4	1.7	2.5	2.1	1.2	1.6
314	C	pla	IM	7394	3	dead	12	10/14/2010	1	1	6.9	1.6	2.0	8.5	2.8
328	C	pla	IM	4759	4	dead	10	10/12/2010	2.8	3.2	5.9	4.8	0.5	9.3	2.4
216	D	pla	SC	4775	1	dead	12	10/14/2010	3.5	2.3	13.3	2.3	2.2	14.2	2.4
283	D	pla	SC	4775	5	live	14	10/16/2010	3	2	6.4	3.4	1.2	10.8	1.9
214	D	pla	SC	4775	7	dead	12	10/14/2010	3.2	2.7	21.4	7.7	9.5	23.3	3.2

(b) Wrong

Figure 1: Common formatting errors that delay analysis (b) and the correct formatting for the same data (a)