

LexisNexis Appeals Data Review

May 12, 2020

Tom Balmat, Duke University

To date, two primary data sets have been received from LexisNexis (LN): set A (3/13/2019) and set B (12/3/2019). The following summarizes the results of several tests of integrity and comparisons of data sets received.

List of Sections

1	Legal Topics	2
1.1	Topic IDs that do not appear in the master topic table	2
1.2	Difference, between data sets, in the distribution of legal topics assigned, by year and by case	3
1.3	Cases with top differences in legal topics counts between data sets A and B	6
1.4	Possibility of copying legal topics from data set A to data set B	9
2	Citations	10
2.1	Citations with no corresponding case header record	10
2.2	Citation records with extraneous commas	11
3	Case type	12
4	Decision date	13
5	Court	14
6	Publication status	16

1 Legal Topics

1.1 Topic IDs that do not appear in the master topic table

Table 1 lists topic IDs that are assigned to cases, but do not appear in the master topics table. Note that, although frequencies differ, the set of unrecognized IDs is the same for both data sets, with the exception of ID C15FD5D77534430BBE65B9AE681A5282 appearing in data set B only.

Table 1: Data sets A and B legal topic IDs that do not appear in the master topic table

Topic ID	Length	n(Set A)	n(Set B)
024C354AF46A4A2FA751B8986DD10471	32	20	22
1FD8BB9677D9467C9BADEB256362B109	32	1	1
25236D06A3164C7E8E6A590AA2BE6D48	32	23	29
4F449B05E21948F09AF3AF140AAF1B87	32	1	5
82FA08A9928B4CF0893760370FF80E2E	32	1	6
8DDF4DB4FE4C487A97DCEED7E7F9D696	32	1	1
C0BE07E8370E4C76AB1D1976E848D4F2	32	3	25
C15FD5D77534430BBE65B9AE681A5282	32	0	2
DDF6E78E4E724C78865E235054136851	32	3	15
DFBA247750FF42F4AA2AC464EBCCE7DE	32	1	3

1.2 Difference, between data sets, in the distribution of legal topics assigned, by year and by case

Legal topic identifiers appearing in the source data are global unique identifier (GUID) codes assigned by LN.¹ It is assumed that each topic has one and only one GUID. Legal topic GUIDs appear in the *legaltopics* column of the supplied case header records and, since a case may have multiple legal topics assigned, the *legaltopics* column concatenates all topic codes and delimits them with a vertical bar symbol. Table 2 contains example *legaltopics* entries from data set A. Text definitions corresponding to legal topic GUIDs were supplied by LN in a separate file. To date, a single definition file has been supplied. To evaluate legal topics, individual topic GUIDs were parsed from *legaltopics* entries using the bar symbol as delimiter, then stored in individual records along with corresponding LN case identifiers (LNI code).

Figure 1 shows the distribution, by year, of the total number of topics assigned to all cases. There is an apparent significant difference in the number of topics assigned between data sets A and B for years 1974 through 2004. Since legal topics are taken directly from supplied case header records, it is assumed that LN’s methods of combining legal topics with case identifiers were very different for the two data sets. Figure 2 shows the distribution of cases by the number of topics assigned. The x-axis indicates number of topics assigned to a case and the y-axis indicates the number of cases with the corresponding topic count. There is an apparent difference in assignment frequency, with cases in data set A (solid line) generally having more topics assigned than those in data set B (dashed line).

Table 2: Example legal topics entries from data set A

Bar delimited topic GUIDs
2956B9EA3A27482F8DC7C512EEB3382E 6C27689629DD41E0BB3EED9958E286F5 FE67EB8441EB4FAAA9AAB2E72461EC0E A80A6FC2A83A47ADBB0CB3B1A427C1D3 98EBABE67ED14D2C8B18FDBC0931E848 FF036325A86142D688DB44AD1244F50E EFF43A3CDC7742D3BA0237DE071A8596 FE67EB8441EB4FAAA9AAB2E72461EC0E D743191585F441CDA8BAA76058C1E979 87944DEFC4E346CFAB58C74C700436D6 FEB27F5D59874496B283A0CBD785A52A 176FEA47A5374CB38D0121E2B4F8CB1B 91A0AEF6DA4E43249D72C1A1142301BE FE67EB8441EB4FAAA9AAB2E72461EC0E D743191585F441CDA8BAA76058C1E979
D1C1B86FA57C4F4DBB80AD6B4C253045 0DFE2750D1294DBA936B6AAECB79C77F 53C6439F34D248E4802F52A3FF0A2C69 CE5CB84ADC5B454F876501D674E8EE03 63E2184C786D48D78C64C10123E84513 63A4F3BF48684B36A1377A48125858E8 2D31918C5FEA4EE3B0B8615766AB64D8 6F9C7A60160B4AC4BCE3D5B340FC6359 D912E7339F584726927A6E2EC45569F5 CB72E53D028A410F8CB4105428B699DF
BF3E916BEE5B43328B9E8B83FE7BFFFD 263745897C814460B589FEE34A3BC1FF 95254DF44FC44686890E9E60C88F3524 DBEE2FC1016747C18110F1EE66EB9085 53E35D0D8C924C49B4EDCE2E79374F84 74062A1EAD004EE5B06143BECB069905 36B57BCD694C46198616654DAEF60AC1 C1BE61018B3B4064BE67707D23B03791 86C67C1B167F4DCFAC700E2848B50F59 08F8A06AD58C4D21B6A3C223DFEE1788

¹A GUID is a 16 byte, computer generated, binary value that, when generated, is guaranteed to be unique from all other codes that were or will be assigned by any request, either from within or outside the organization.

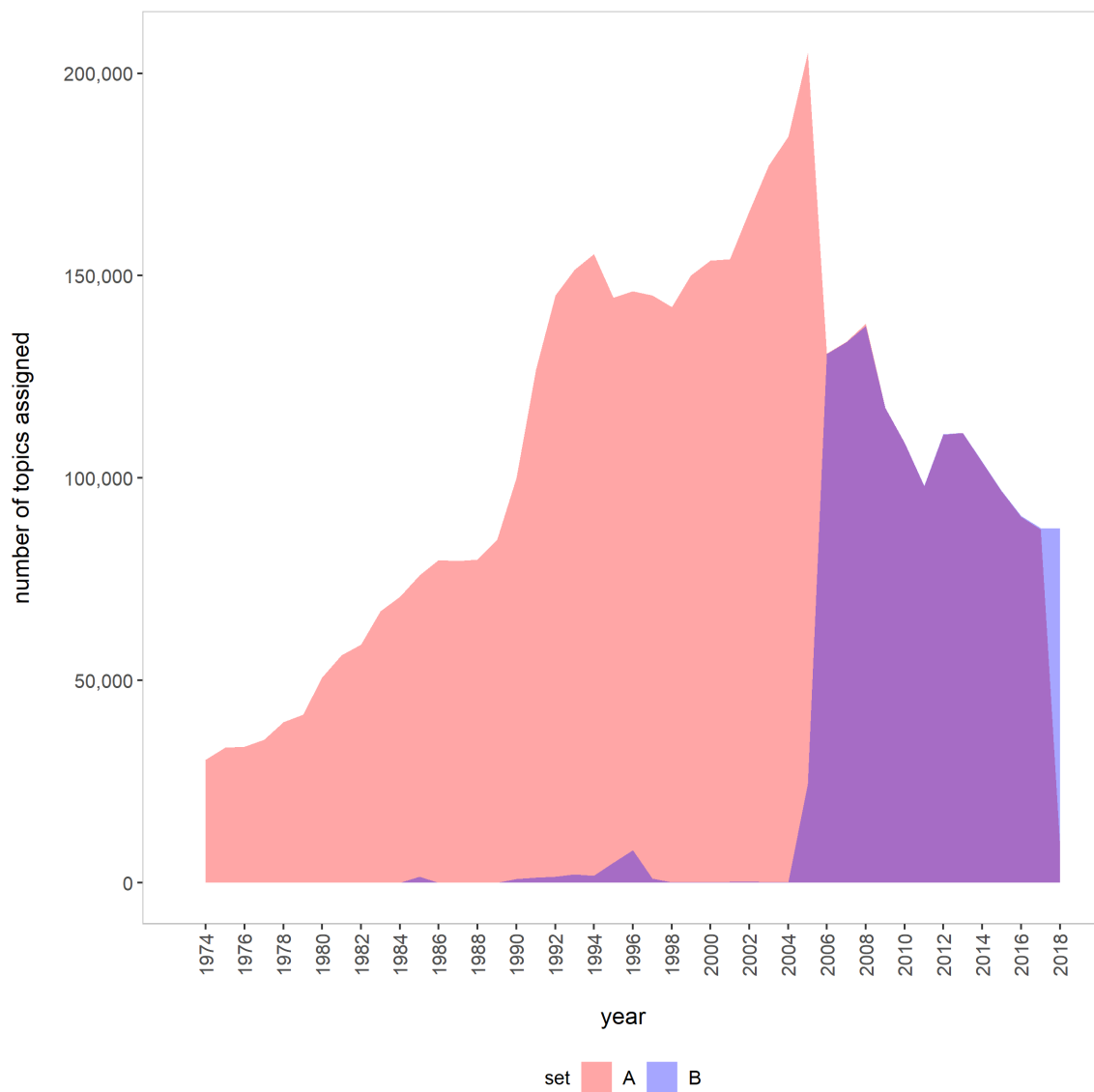


Figure 1: Distribution of total number of topics assigned by data set and year

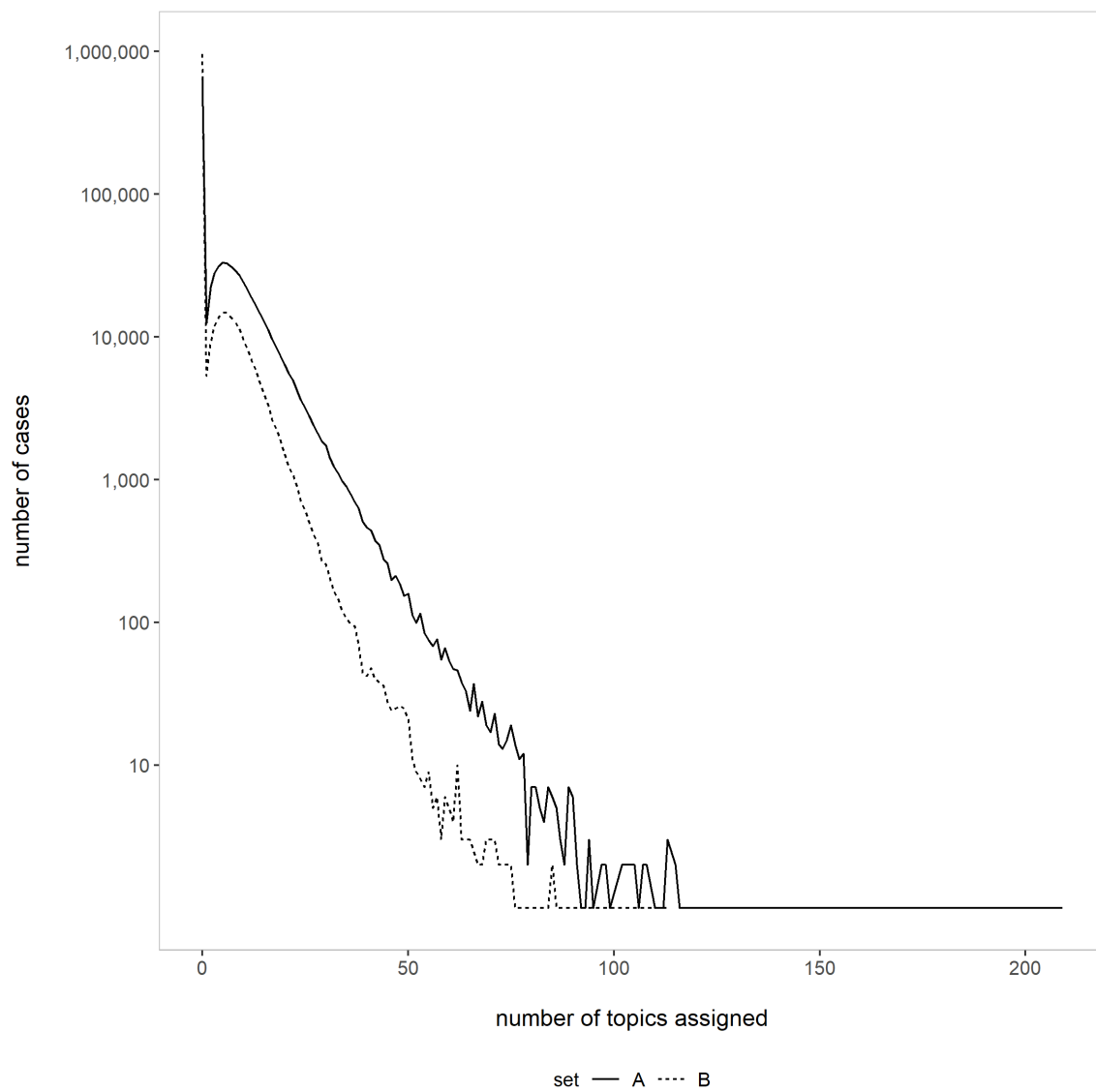


Figure 2: Distribution of cases by data set and number of topics assigned

1.3 Cases with top differences in legal topics counts between data sets A and B

Table 3 lists the top 100 cases in descending order of difference in legal topics assigned. Note that all cases in the table appear in both data sets A and B.

Table 3: Top 100 cases with a difference in legal topics assigned in data sets A and B

Title	Court	Date	n(set A)	n(set B)
United States v. Yousef	2nd Circuit	2003-04-04	209	0
United States v. Salameh	2nd Circuit	1998-08-04	141	0
Data Gen. Corp. v. Grumman Sys. Support Corp.	1st Circuit	1994-09-14	137	0
United States v. Phillips	5th Circuit	1981-12-28	133	0
United States v. Soto-Beniquez	1st Circuit	2003-11-20	129	0
United States v. Darden	8th Circuit	1995-11-22	124	0
United Int'l Holdings, Inc. v. Wharf (Holdings) Ltd.	10th Circuit	2000-04-28	123	0
United States v. Allen	8th Circuit	2001-04-12	119	0
United States v. Schlei	11th Circuit	1997-09-18	118	0
United States v. McVeigh	10th Circuit	1998-09-08	116	0
United States v. Hammoud	4th Circuit	2004-09-08	115	0
United States v. Haldeman	DC Circuit	1976-10-12	115	0
United States v. Higgs	4th Circuit	2003-12-22	113	0
United States v. Nelson-Rodriguez	1st Circuit	2003-02-07	113	0
United States v. Hall	5th Circuit	1998-08-21	112	0
Coalition for Gov't Procurement v. Fed. Prison Indus.	6th Circuit	2004-04-12	110	0
United States v. Moya-Gomez	7th Circuit	1988-09-30	108	0
United States v. Webster	5th Circuit	1998-12-03	108	0
United States v. Solis	5th Circuit	2002-07-18	107	0
Unitherm Food Sys. v. Swift-Eckrich, Inc.	Federal Circuit	2004-07-12	107	0
United States v. Cope	6th Circuit	2002-11-19	106	0
Gall v. Parker	6th Circuit	2000-10-30	105	0
United States v. N.	DC Circuit	1990-07-20	105	0
United States v. Voigt	3rd Circuit	1996-07-09	102	0
United States v. Collazo-Aponte	1st Circuit	2000-06-27	102	0
Pennzoil Co. v. Federal Energy Regulatory Com.	5th Circuit	1981-05-20	99	0
United States v. Gaskin	2nd Circuit	2004-04-16	98	0
United States v. Jackson	10th Circuit	2000-06-02	98	0
Byrd v. Collins	6th Circuit	2000-04-06	97	0
United States v. Thomas	5th Circuit	1993-12-21	97	0
Elliott Indus. v. BP Am. Prod. Co.	10th Circuit	2005-05-10	95	0
United States v. Sepulveda	1st Circuit	1993-12-20	94	0
United States v. Yarbrough	9th Circuit	1988-07-06	94	0
United States v. Tarantino	DC Circuit	1988-04-12	93	0
Mass. Eye & Ear Infirmary v. QLT Phototherapeutics, Inc.	1st Circuit	2005-06-16	92	0
Hobson v. Wilson	DC Circuit	1984-06-08	91	0
United States v. Graham	6th Circuit	2001-12-17	91	0
Williams v. Bagley	6th Circuit	2004-08-13	90	0
United States v. Diaz	2nd Circuit	1999-05-04	90	0
United States v. Morrow	5th Circuit	1999-05-25	90	0
United States v. Fernandez	9th Circuit	2004-10-27	90	0
United States v. Shryock	9th Circuit	2003-09-04	90	0

Table 3: Top 100 cases with a difference in legal topics assigned in data sets A and B

Title	Court	Date	n(set A)	n(set B)
Richardson v. Reno	11th Circuit	1998-12-22	90	0
United States v. Blakeney	6th Circuit	1991-08-23	89	0
Buell v. Mitchell	6th Circuit	2001-12-04	89	0
Hale v. Gibson	10th Circuit	2000-09-25	89	0
United States v. Mann	5th Circuit	1998-11-23	89	0
Bowers v. NCAA	3rd Circuit	2003-08-20	89	0
Williams v. Woodford	9th Circuit	2002-09-10	89	0
Bell v. Milwaukee	7th Circuit	1984-09-04	89	0
United States v. Gibbs	6th Circuit	1999-04-16	88	0
Hemmings v. Tidyman’s Inc.	9th Circuit	2002-04-11	88	0
United States v. Novaton	11th Circuit	2001-10-30	87	0
Cable/Home Commun. Corp. v. Network Prods.	11th Circuit	1990-06-04	87	0
United States v. Tolliver	5th Circuit	1995-08-14	87	0
United States v. Miller	2nd Circuit	1997-06-20	86	0
United States v. Aguilar	9th Circuit	1989-03-30	86	0
United States v. McGlory	3rd Circuit	1992-06-19	86	0
Shotz v. City of Plantation	11th Circuit	2003-09-08	86	0
United States v. Daychild	9th Circuit	2004-02-17	85	0
United States v. Souffront	7th Circuit	2003-08-06	85	0
Lightfoot v. Union Carbide Corp.	2nd Circuit	1997-03-27	85	0
United States v. Sturman	6th Circuit	1991-10-24	85	0
Streber v. Hunter	5th Circuit	2000-08-04	84	0
United States v. Thomas	5th Circuit	1994-01-25	84	0
MASON v. MITCHELL	6th Circuit	2003-02-06	84	0
Coleman v. Mitchell	6th Circuit	2001-10-10	84	0
Ford Motor Co. v. Summit Motor Prods.	3rd Circuit	1991-04-08	84	0
United States v. Pungitore	3rd Circuit	1990-08-01	84	0
Rose v. Bartle	3rd Circuit	1989-03-20	83	0
Cox v. Administrator United States Steel & Carnegie	11th Circuit	1994-04-05	83	0
Retirement Fund Trust of Plumbing v. Franchise Tax Bd.	9th Circuit	1990-07-16	83	0
Atlantic Richfield Co. v. Farm Credit Bank	10th Circuit	2000-09-13	82	0
Coosewoon v. Meridian Oil Co.	10th Circuit	1994-05-25	82	0
In re Exxon Valdez	9th Circuit	2001-11-07	82	0
United States v. Mansoori	7th Circuit	2002-08-29	82	0
United States v. Logan	6th Circuit	1999-07-19	82	0
United States v. Doerr	7th Circuit	1989-10-03	81	0
American-Arab Anti-Discrimination Comm. v. Reno	9th Circuit	1995-11-08	81	0
United States v. Baker	9th Circuit	1993-10-19	81	0
United States v. Bakhshkooei	9th Circuit	1997-03-19	81	0
United States v. Moore	7th Circuit	1991-07-19	81	0
United States v. Weiner	9th Circuit	1978-05-15	81	0
Transgo, Inc. v. Ajac Transmission Parts Corp.	9th Circuit	1985-01-15	81	0
In re Cmty. Bank of N. Va. & Guar. Nat’l Bank of Tallahassee Second Mortg. Loan Litig.	3rd Circuit	2005-08-11	80	0
Koslow v. Pennsylvania	3rd Circuit	2002-08-21	80	0
United States v. Peters	7th Circuit	1986-05-20	80	0
United States v. Arias-Villanueva	9th Circuit	1993-07-20	80	0

Table 3: Top 100 cases with a difference in legal topics assigned in data sets A and B

Title	Court	Date	n(set A)	n(set B)
Image Tech. Servs. v. Eastman Kodak Co.	9th Circuit	1997-08-26	80	0
United States v. Chandler	11th Circuit	1993-07-19	80	0
T. J. Stevenson & Co. v. 81,193 Bags of Flour	5th Circuit	1980-10-27	79	0
Permian Petroleum Co. v. Petroleos Mexicanos	5th Circuit	1991-06-27	78	0
Loftin & Woodard, Inc. v. United States	5th Circuit	1978-08-09	78	0
United States v. Griffin	5th Circuit	2003-03-10	78	0
United States v. Garcia	9th Circuit	1994-02-15	78	0
United States v. Briscoe	7th Circuit	1990-02-26	78	0
Chavez v. Thomas & Betts Corp.	10th Circuit	2005-01-24	78	0
United States v. Pedraza	10th Circuit	1994-06-30	78	0
Carcieri v. Norton	1st Circuit	2005-09-13	78	0
United States v. Ramirez-Lopez	9th Circuit	2003-01-10	78	0

1.4 Possibility of copying legal topics from data set A to data set B

Many cases appearing in data set B with no legal topics appear in data set A with legal topics. On the assumption that topics were correctly assigned in data set A and LNI assignment by case is consistent between data sets, one option is to copy topics from data set A to data set B, maintaining LNI assignment from data set A. The number of topics assigned in data set B is 1,463,602 or 31.2% of the 4,699,274 assigned in data set A. Figure 1 indicates that most topics assigned between years 1974 and 2004 in data set A are absent in data set B, but suggests consistent frequency of topics by year between 2005 and 2017. Examining cases with topics in both data sets reveals that 99.3% of topics assigned to these cases appear in both data sets. This suggests that, when topics were assigned in data set B, the method of assignment was consistent with the method used in data set A. Table 4 shows the distribution, by year, of legal topics assigned in data set B that do not appear in data set A. It is seen that the majority of these cases appear in 2016, 2017, and 2018, years for which cases were added in data set B. Of the 4,699,176 topics assigned in data set A with corresponding case in data set B, 1,383,618 (29.4%) also appear in data set B. Consequently, if legal topics are copied from data set A to data set B, over 70% of topics in data set B will have been introduced by copying, as opposed to having been provided directly by LN.

Table 4: Distribution of topics appearing in data set B that do not appear in data set A

Year	Number of topics
1973	36
1982	1
1985	8
1987	3
1992	10
1997	5
1999	13
2002	8
2007	13
2008	22
2009	8
2010	19
2011	33
2012	53
2013	41
2014	48
2015	66
2016	207
2017	395
2018	78,995

2 Citations

2.1 Citations with no corresponding case header record

Citations were received with data set A only. When asked whether additional or revised citations would be included with data set B, LN stated that no changes were made to the citations file as a result of the extract definition of data set B. Twenty five LNIs appear in the citations file as a citing case, but with no corresponding case header record in data set B.² Table 5 lists these LNIs.

Table 5: Citing LNIs that do not appear in case header records of data set B

LNI
3RFY-1PP0-0038-X128-00000-00
3S4N-9710-0039-S2RH-00000-00
3S4N-SYC0-0054-74RK-00000-00
3S4X-4NR0-008H-V18P-00000-00
3S4X-62M0-003B-P4B0-00000-00
3S4X-7T40-001B-K49W-00000-00
3S4X-DM40-003B-52S8-00000-00
3S4X-DM60-003B-52SY-00000-00
3S4X-DMD0-003B-52VD-00000-00
3S4X-F5V0-003B-P4NK-00000-00
3S4X-J840-003B-P2MK-00000-00
3T3Y-TS60-0038-X1MJ-00000-00
3T8C-CSW0-0038-X3XN-00000-00
3TP6-S860-0038-X120-00000-00
41GG-8X00-0038-X4N9-00000-00
486S-HBX0-0038-X0VX-00000-00
4FN6-TKH0-0038-X19G-00000-00
4VW5-8J60-TXFN-62T3-00000-00
582C-S0Y1-F04K-N1PH-00000-00
583F-SJC1-F04K-W0KY-00000-00
583G-RMS1-F04K-N1VC-00000-00
58NX-0P41-F04K-V001-00000-00
5BXT-9H71-F04K-N000-00000-00
5KSG-TH21-DYXG-1002-00000-00
5RYJ-7CK1-F1WF-M1SS-00000-00

²11,799 citing LNIs are absent in the case header records of data set A

2.2 Citation records with extraneous commas

133 records have citation entries that contain one or more commas which, in a CSV context, cause additional columns to be inferred during import and generates an error. Example records with commas in the citation field appear in table 6. Some records have commas surrounded by double quotes, as in the final record. This is also an invalid CSV format.

Table 6: Example citation records with comma in citation field

Citing LNI	Cited LNI	Citation	Shepard's treatment
3S4X-4310-0039-M2XJ-00000-00	3S4X-0X80-0039-X1TB-00000-00	Md. Code Ann., Health Occ. sec. 7-101 (h)	f
3TC7-NDR0-0038-X0BH-00000-00	3S4X-HDB0-003B-P4ST-00000-00	Wash. Cty., Md. Pub. Loc. L. 1930 sec. 572 (1941C526)(i)	f
3S4X-J2P0-003B-P104-00000-00	3S4X-H480-008H-V1HX-00000-00	Copyrights Class TX Minnesota Multiphasic Personality Schedule, The	f
3S4X-17P0-0039-M00R-00000-00	3S4X-2490-0039-M0J0-00000-00	Worc. Cty., Md. Pub. Loc. L. 1930 sec. 193	d
3S4X-2J60-001B-K469-00000-00	3S4X-3F10-003B-G4S4-00000-00	Kan. Stat. Ann. sec. 66-1,112g	f
3S4W-YSY0-0039-P47G-00000-00	3S4X-1RS0-003B-G2KJ-00000-00	Neb. Rev. Stat. sec. 81-8,110.13	d
3S4X-35G0-003B-52YN-00000-00	3S4X-JNV0-0039-P02C-00000-00	Balt. Cty., Md. Cty. Code 1978 2 (22-74)	f
3S4X-5RB0-0039-W25N-00000-00	3S4W-VGK0-0039-M08C-00000-00	Trademarks Bandido by Cervantes, El	d
3S4W-YSH0-006F-M3TY-00000-00	3S4X-2230-008H-V2NH-00000-00	Md. Code Ann., Health Occ. sec. 14-316 (c)	f
3S4X-77J0-003B-53H9-00000-00	3S4W-WHD0-0039-M55G-00000-00	Trademarks Dabs Creation, A	d
3S4X-0B30-0039-M0T3-00000-00	3S4X-15M0-0039-Y0SX-00000-00	Md. Code Ann., Fin. Inst. sec. 6-215 (d)	e
3S4X-5HG0-003B-50JP-00000-00	3S4V-S420-0054-8054-00000-00	Copyrights Class PA Jungle Killer, The	f
3S4X-5950-0039-W1N8-00000-00	3S4V-P6H0-003B-20JV-00000-00	Wicomico Cty., Md. Pub. Loc. L. 1959 sec. 23E	f
3S4X-1NT0-003B-P4CG-00000-00	3S4X-HDB0-003B-P4ST-00000-00	Wash. Cty., Md. Pub. Loc. L. 1930 sec. 572 (1941C526)(i)	f
3S4X-2BN0-003B-P2W7-00000-00	3S4X-5V90-003B-P34Y-00000-00	Balt. Cty., Md. Pub. Loc. L. 1948 sec. 399	d
3S4X-KJS0-0039-W0B6-00000-00	3S4W-XP90-0039-M4CC-00000-00	Pr. Geo. Cty., Md. Pub. Loc. L. 1953 sec. 1117	d
4401-R1M0-0038-X1R7-00000-00	3S4X-5V90-003B-P34Y-00000-00	Balt. Cty., Md. Pub. Loc. L. 1948 sec. 399	ca
3S4X-B4R0-003B-507G-00000-00	3S4X-HRM0-0039-P1Y4-00000-00	Harford Cty., Md. Pub. Loc. L. 1965 sec. 214	d
3S4X-40K0-003B-G262-00000-00	3S4X-0X80-0039-X1TB-00000-00	Md. Code Ann., Health Occ. sec. 7-101 (h)	f
3S4W-T9N0-0039-M0FN-00000-00	3S4W-WHD0-0039-M55G-00000-00	Trademarks Dabs Creation, A	f
3S4X-2NV0-006F-M138-00000-00	3S4X-8XW0-0039-P2SX-00000-00	Trademarks Better Jacket-Thru Science“, ”A	d

3 Case type

Cases in the source data are classified as either civil, criminal, civil and criminal, or are not assigned a type. A small number of cases have classifications that differ between data sets A and B. These appear in table 7.

Table 7: Cases with different classifications in data sets A and B

LNI	Civil		Criminal	
	Set A	Set B	set A	set B
3S4X-9FW0-001B-K3G7-00000-00	no	yes	no	no
3S4X-9GP0-003B-51GH-00000-00	no	yes	no	no
3S4X-G600-003B-P1JH-00000-00	no	yes	no	no
4076-FFT0-0038-X2JM-00000-00	no	yes	no	yes
4239-4R30-0038-X3TJ-00000-00	no	yes	no	no
4344-WKG0-0038-X04X-00000-00	no	yes	no	yes
43CB-G040-0038-X4NG-00000-00	no	yes	no	no
43XX-3MG0-0038-X0YJ-00000-00	no	yes	no	no
44BG-8ST0-0038-X0F2-00000-00	no	yes	no	no
45V7-KX00-0038-X2R7-00000-00	no	yes	no	no
469G-P1T0-0038-X485-00000-00	no	yes	no	no
46RD-5520-0038-X3S7-00000-00	no	yes	no	yes
46VM-M2T0-0038-X0T3-00000-00	no	yes	no	no
47TG-1WB0-0038-X10M-00000-00	no	yes	no	no
48WJ-V5N0-0038-X2BG-00000-00	no	no	no	yes
4CGP-WYW0-0038-X1K2-00000-00	no	yes	no	no
4CT5-XYH0-0038-X09G-00000-00	no	no	no	yes
4F80-5950-0038-X3JB-00000-00	no	yes	no	no
4GC3-XC80-0038-X2BT-00000-00	no	yes	no	no
4GM4-Y5S0-0038-X4WH-00000-00	no	yes	no	no

4 Decision date

Valid decision dates in the source data have format yyyy-mm-dd. Of all cases appearing in both data sets A and B, one had differing decision dates between sets: LNI 3S4W-X6F0-0039-M2H4-00000-00, with dates 1978-01-01 (in set A) and 1978-11-07 (in set B).

5 Court

Table 8 lists LNIs for cases with an empty court name. Table 9 lists cases that appear in both data sets A and B, but with different court assignments.

Table 8: Cases with an empty court name in source records

LNI	Exists in set A	Exists in set B
3RHM-PKM0-0038-X3K4-00000-00	yes	no
3RV5-KP00-008H-V07P-00000-00	yes	yes
3RV5-KSB0-003B-50M4-00000-00	yes	yes
3S12-VYB0-0038-X3BN-00000-00	yes	no
3S4W-T920-0039-M08N-00000-00	yes	yes
3S4W-WSK0-0039-M0TG-00000-00	yes	no
3S4X-0650-0039-W4RT-00000-00	yes	no
3S4X-1PS0-001B-K1XV-00000-00	yes	yes
3S4X-5JR0-003B-50T8-00000-00	yes	yes
3S4X-6M00-008H-V0GV-00000-00	yes	no
3S4X-6M00-008H-V0GW-00000-00	yes	no
3S4X-BYM0-001B-K3HW-00000-00	yes	yes
3S4X-C0G0-001T-D46N-00000-00	yes	no
3S4X-CSF0-003B-5370-00000-00	yes	yes
4F7Y-YXX0-TVRV-K31B-00000-00	yes	no
4G24-T8Y0-TVRV-K2NF-00000-00	yes	no
4KND-FMG0-TVRV-K3DG-00000-00	yes	yes
5CK5-5WW1-F04K-M22K-00000-00	no	yes
5NMN-RCM1-DYTB-X1V5-00000-00	yes	no
5NMN-RCT1-DYTB-X1V9-00000-00	yes	no
5NMN-RCT1-DYTB-X1VC-00000-00	yes	no
5NMN-RCY1-DYTB-X1VF-00000-00	yes	no
5NMN-RD71-JCB9-61DG-00000-00	yes	no
5NMN-RD81-JCB9-61DM-00000-00	yes	no
5NMN-RD91-JCB9-61DP-00000-00	yes	no
5NMN-RDC1-JCB9-61DX-00000-00	yes	no
5NMN-RDF1-JCB9-61F0-00000-00	yes	no
5NMN-RDN1-JCB9-61F1-00000-00	yes	no
5NMN-RDR1-JCB9-61F4-00000-00	yes	no
5NMN-RFF1-DYTB-X1VK-00000-00	yes	no
5NMN-RGW1-DYTB-X1VV-00000-00	yes	no
5NMN-RH11-JCB9-61FX-00000-00	yes	no
5NMN-RH41-JCB9-61G4-00000-00	yes	no
5NMN-RJ11-JCB9-61GD-00000-00	yes	no
5NMN-RJ41-JCB9-61GJ-00000-00	yes	no
5NRR-WCV1-F04K-W0B6-00000-00	yes	no
5NYK-PX51-DYTB-X4WP-00000-00	yes	no
5NYK-PXX1-DYTB-X4XM-00000-00	yes	no
5NYK-PXX1-DYTB-X4XP-00000-00	yes	no
5NYK-PY01-DYTB-X4XR-00000-00	yes	no
5RJ8-3DP1-F04K-S2GK-00000-00	yes	no
5RK7-J161-DYFH-X0W8-00000-00	yes	no
5V50-0W61-JGBH-B2TG-00000-00	no	yes
8R3J-55T2-DHRX-K007-00000-00	no	yes

Table 9: Cases appearing in data sets A and B, but with different court names

LNI	Court set A	Court set B
3S4W-WSK0-0039-M0TG-00000-00	<i>empty</i>	5th Circuit Court of Appeals
58YB-N3S1-F04K-G01C-00000-00	11th Circuit Court of Appeals	5th Circuit Court of Appeals
5RK7-J161-DYFH-X0W8-00000-00	<i>empty</i>	5th Circuit Court of Appeals
5RJ8-N4C1-F361-M16P-00000-00	7th Circuit Court of Appeals	11th Circuit Court of Appeals
5RF1-H0G1-FJDY-X26K-00000-00	9th Circuit Court of Appeals	6th Circuit Court of Appeals
5NRR-WCV1-F04K-W0B6-00000-00	<i>empty</i>	10th Circuit Court of Appeals
3S4X-0650-0039-W4RT-00000-00	<i>empty</i>	3rd Circuit Court of Appeals
5PTK-GP41-F2TK-23SR-00000-00	5th Circuit Court of Appeals	1st Circuit Court of Appeals
5RG5-6JJ1-JBT7-X08T-00000-00	5th Circuit Court of Appeals	1st Circuit Court of Appeals
3RV5-F210-0039-P06K-00000-00	Federal Circuit Court of Appeals	Temporary Emergency Court of Appeals
3S4W-V570-0039-X2GT-00000-00	Federal Circuit Court of Appeals	Temporary Emergency Court of Appeals
3S4W-VBF0-0039-M4VW-00000-00	Federal Circuit Court of Appeals	Temporary Emergency Court of Appeals
3S4W-X0C0-0039-X410-00000-00	Federal Circuit Court of Appeals	Temporary Emergency Court of Appeals
3S4W-XB90-0039-X4GW-00000-00	Federal Circuit Court of Appeals	Temporary Emergency Court of Appeals
3S4X-12X0-0039-M49S-00000-00	Federal Circuit Court of Appeals	Temporary Emergency Court of Appeals
3S4X-12X0-0039-M49T-00000-00	Federal Circuit Court of Appeals	Temporary Emergency Court of Appeals
3S4X-2J90-0039-M45P-00000-00	Federal Circuit Court of Appeals	Temporary Emergency Court of Appeals
3S4X-2WJ0-0039-M0N5-00000-00	Federal Circuit Court of Appeals	Temporary Emergency Court of Appeals
3S4X-68W0-0039-W32D-00000-00	Federal Circuit Court of Appeals	Temporary Emergency Court of Appeals
3S4X-K050-0039-W4FW-00000-00	Federal Circuit Court of Appeals	Temporary Emergency Court of Appeals
3S4X-KMP0-0039-W0FK-00000-00	Federal Circuit Court of Appeals	Temporary Emergency Court of Appeals
5CK5-5WW1-F04K-M22K-00000-00	5th Circuit Court of Appeals	<i>empty</i>
56F6-GPM1-F04K-P21S-00000-00	9th Circuit Court of Appeals	Judicial Conference, Committee on Judicial Conduct
5BWH-9G81-F04K-J00J-00000-00	9th Circuit Court of Appeals	Judicial Conference, Committee on Judicial Conduct

6 Publication status

Each case header record has publication status of either “Reported” or “Unreported.” Table 10 lists the number of cases, by year of decision date, that have different publication status values in the two data sets. Note that cases with decision dates in years 1974, 1976, 1979, and 1981 have consistent status values in both data sets. Data set A contains no cases with a decision date in 1973.

Table 10: Number of cases with different publication status values in data sets A and B

Year	n(set A rep \rightarrow set B unrep)	n(set A unrep \rightarrow set B rep)
1975	0	1
1977	0	1
1978	0	1
1980	0	1
1982	0	7
1983	0	7
1984	0	4
1985	0	8
1986	0	4
1987	0	6
1988	0	16
1989	0	15
1990	0	10
1991	0	6
1992	0	26
1993	2	18
1994	0	24
1995	0	15
1996	0	19
1997	1	46
1998	1	47
1999	0	21
2000	0	31
2001	0	9
2002	0	10
2003	0	11
2004	0	6
2005	0	9
2006	0	10
2007	0	15
2008	0	7
2009	0	11
2010	0	8
2011	0	8
2012	4	20
2013	3	20
2014	3	47
2015	6	28
2016	5	27
2017	2	58
2018	0	5