OBTS Data Inventory

Introduction

OBTS Main Database

The Offender Based Transaction Statistics (OBTS) database is designed to provide information about trends in how felony arrests are handled in California. OBTS is an extract of California's Automated Criminal History System (ACHS), which is used, among other things, for background checks. The OBTS captures the arrest-cycle for all adult-level felony arrests (includes juveniles treated as adults), beginning with arrest itself and ending with a release, a decline to prosecute, a not guilty verdict, or a guilty verdict and a sentence. OBTS data begin in 1982 and contain 200,000 to 350,000 records per year. Each record includes the arresting officer's jurisdiction, the arrest date and BCS code for offense, the age/birth date of the arrestee, the arrestee's demographics, and, if applicable, charges pressed, charges convicted, and sentence type. The OBTS also contains direct identifiers such as name, CII, and DoB.

Background:

- Only 65-75% of felony arrest dispositions are captured annually (DO WE KNOW WHY?)
- The representativeness of OBTS data may vary by county (CONFIRM WITH MARIE OR RANDIE)
- Disposition data come from form link, which is filled out by law enforcement, the district attorney's office, and the courts
- OBTS only captures final dispositions, intermediate dispositions such as diversion programs or retrials are not included [WHAT DEFINES A DISPOSITION AS FINAL?]
- If there are multiple arrest offenses, the one shown is the most serious offense in terms of potential incarceration sentence
- If there are multiple conviction offenses, the one shown is the most serious offense in terms of potential incarceration sentence
- OBTS contains data about juveniles whose final disposition was in the adult criminal justice system

Historical Changes:

- Beginning in 1998, lower and superior courts could unify their data reporting
- From 1992 through 1997, convictions were prioritized for data entry, non-conviction records may be missing in higher numbers
- [FILE CREATION DATES AFFECTS NUMBER OF RECORDS PER YEAR - HOW DOES THIS WORK?]
- Offense data for 2002 may be missing due to data being entered in the wrong field

County-specific issues:

- Alameda: under-reporting of final dispositions from 1998 through 2002
- Contra Costa: Incomplete 2008 data
- Los Angeles: Missing thousands of dispositions from 1998 through 2000
- Sacramento: Incomplete 2008 data
- San Bernardino: Incomplete data from 1999 through 2001
- San Diego: Decrease in dispositions after 2000 may be result of a reporting issue
- San Francisco: Law enforcement makes very few final dispositions, more likely to send to DA; 2004 data may have errors; 2005 and 2010 data are incomplete;
- Santa Barbara: Data for 1993 through 1997 may be unreliable
- Ventura: Incomplete data from 1996 through 2001; incomplete data for 2010

Sample Rows

id	cii_record_ty...	cii_number	pdr_record_id	name	birth_year	...
1	OBN	SEALED	unknown	SEALED ...	1993	...
2	automated	SEALED	current_offen...	SEALED ...	1932	...
3	automated	SEALED	current_offen...	SEALED ...	1933	...
4	automated	SEALED	current_offen...	SEALED ...	1931	...
5	automated	SEALED	current_offen...	SEALED ...	1932	...
...	...	...	...	...	...	...

Variable Summary

name	type	value	description
id	numeric	1-9654994	
cii_record_type	factor	automat.../hybrid/m...	
cii_number	character	pii	Criminal Identification and Information number
pdr_record_id	factor	current.../new_off...	Personal Descriptor Record ID
name	character	pii	
birth_year	integer	1891-1999	
birth_month	integer	1-12	
birth_day	integer	1-31	
gender	factor	female/male/unknown	
race	factor	america.../asian_i....	
deceased	factor	deceased/decease.....	whether individual is deceased
arrest_record_id	factor	final_c.../police/p...	last reported step in crim justice process
arrest_bypass	factor	no_bypa.../not_cor....	
arrest_converted_data	factor	convert.../not_con...	converted from old format
arrest_event_year	integer	1929-2015	
arrest_event_month	integer	1-12	
arrest_event_day	integer	1-31	
arresting_agency	factor	010000/010001/01010.	arresting or booking agency
local_file_number	character	pii	
arrest_charge_type	factor	felony/misdeme...	
multiple_arrest_charges	factor	1/2/3/4/5/6/7/8/9	
arrest_disposition_year	integer	1944-2015	date of arresting officer disposition
arrest_disposition_month	integer	1-12	
arrest_disposition_day	integer	1-31	
arrest_disposition_type	factor	1/2/3/4/5/6/7/10/11...	
arrest_multiple_dispositions	factor	0/1/2/3/4	
prior_record_code	factor	length_.../misc_pr....	
arrest_offense	factor	1004/1005/3005/9004...	
arrest_qualifier	character	.../AQ	
arrest_summary_code	factor	1/2/3/4/5/6/7/8/9/1.	BCS summary code for arrest offense
arrest_edit_error_code	character	/C /B /C...	
court_record_id	factor	lower_c.../superio...	court of final disposition
court_bypass	factor	correct.../no_bypa....	
court_converted_data	factor	1	
court_event_year	integer	1982-2014	last date of court proceedings
court_event_month	integer	1-12	
court_event_day	integer	1-31	
court_judicial_district	factor	01****/01100J/01120...	
court_number	character	pii	

name	type	value	description
court_disposition_type	factor	acquited/certifi......	outcome of court trial
court_multiple_dispositions	factor	1/2/3/4/5/6/7/9	
court_proceeding_type	factor	certifi.../felony_....	
sentence_type	factor	1/2/3/4/5/6	
sentence	factor	CRC/CYA/death/fine/...	
court_charge_type	factor	felony/infract.../m...offense level of conviction crime	
court_multiple_disposition_charges	factor	1/2/3/4/5/6/7/8/9	
court_disposition_offense	factor	1004/3005/4016/8001...	
court_qualifier	factor	.../ KA	
court_summary_code	factor	1/2/3/4/5/6/7/8/9/1...	
court_edit_error_code	character	/D /K /F...	

Judicial Districts

Data were pulled from the 1982-2014 OBTS Tech Manual with the following changes:

- 6 records that did not conform to the practice of `County Name - Location` were modified by hand:

code	old name	county	location
29660J	Nevada County Court Nevada	Nevada	Nevada
31620J	Placer County Court - Colfax- Alta - Dutch Flat	Placer	Colfax/Alta/Dutch Flat
32630J	Plumas County Court - Almanor - Chester Branch	Plumas	Almanor/Chester Branch
34630J	Sacramento Court - Fair Oaks - Folsom	Sacramento	Fair Oaks/Folsom
35640J	San Benito Municip- Hollister	San Benito	Hollister
45633J	Shasta County Court Central Valley	Shasta	Central Valley
55650J	Tuolumne Court _ Fifth (Jamestown)	Tuolumne	Fifth (Jamestown)

- Records for when the word "County" was missing were assumed to be at the county level
- The record with "El Dorada" county was changed to "El Dorado" county
- The one instance of "Superior" court was changed to a county court

Is "Colfax/Alta/Dutch Flat" correct for Placer, as there is already a "Colfax"?

Sample Rows

county_code	agency_code	county	type	agency
01	NA	Alameda	Arrest	No Court Disp...
01	100J	Alameda	County	Alameda Count...
01	120J	Alameda	County	Hayward
01	410J	Alameda	County	Alameda
01	420J	Alameda	County	Berkeley
...	...	...	...	...

Arresting Agencies

Pulled from the 1982-2014 OBTS Tech Manual.

Sample Rows

code	agency	type
010000	OAKLAND	Sheriffs Offi...
010001	OAKLAND	County Jail
010100	ALAMEDA	Police Depart...
01013A	OAKLAND	District Atto...
01013G	OAKLAND	Probation Dep...
...	...	...

CJIS Codes

Pulled from the 1982-2014 OBTS Tech Manual.

Sample Rows

cjis_code	statue_code	code_type	literal	toc	summary_code	hierarchy
48052	20	VC	USE/ETC FALSE...	M	60	141800
48103	27	VC	IMPERSONATE C...	M	60	141600
48051	31	VC	GIVE FALSE IN...	M	60	162600
46039	32	PC	ACCESSORY:PER...	M	60	92600
46004	32	PC	ACCESSORY	F	25	68600
...	...	...	...	...	...	...

Qualifier Codes

Pulled from the 1982-2014 OBTS Tech Manual. Qualifiers form a one-to-many map from records to codes, in which each entry of the OBTS can be associated with more than one qualifier.

Sample Rows

qualifier	literal_short	code_section	code_type	literal_long
AB	ZZZZ	00ATTEMPTED	XX	ATTEMPTED
AC	ZZZZ	00BENCH WARRA...	XX	BENCH WARRANT
AF	ZZZZ	00CONTEMPT	XX	CONTEMPT
AJ	ZZZZ	00FAIL TO APP...	XX	FAIL TO APPEAR
AO	ZZZZ	00PROBATION R...	XX	PROBATION REVO
...	...	...	...	...

Tables

Main Table Variables

cii_record_type

Description

whether the individual in the record has other records that are digitized and whether those records can be linked using fingerprints

Labels

- 1 - automated
- 2- manual
- 3 - OBN (no CII record on file)
- 4 - Hybrid (began 1993)

Notes

automated: fully digitized record; manual: paper records only; hybrid: mix of paper and digitized records; OBN: records that are not linked via fingerprints

Summary

Name	Freq
automated	8782819
OBN	398939
manual	316923
hybrid	100757
NA	55556

cii_number

Description

Statewide unique identification number for an individual.

Prompt

Enter the unique state identification number, referred to as the "CII#", assigned to the subject's criminal history record. If no CII number is available or known to the agency, leave the area blank and continue processing the form.

pdr_id

Description

The Personal Descriptor Record ID indicates whether the offender is new to the system [CORRECT?]

Labels

- 00 - new offender (no arrest before 8/23/82)
- 01 - current offender (arrested before 8/23/82)
- 02 - unknown
- 09 - delete CII

Notes

proportion of current offenders highest when the data begin in 1982 but rapidly decreases to less than 10% - that implies only 10% of people arrested for a felony have a prior felony arrest???

Summary

Name	Freq
new_offender	7238025
current_offender	1701846
unknown	715123

name

Description

Individual's name.

Prompt

Enter the complete name of the subject as it was submitted on the arrest fingerprint submission or as indicated on the citation. Aliases, nicknames, or monikers are not required.
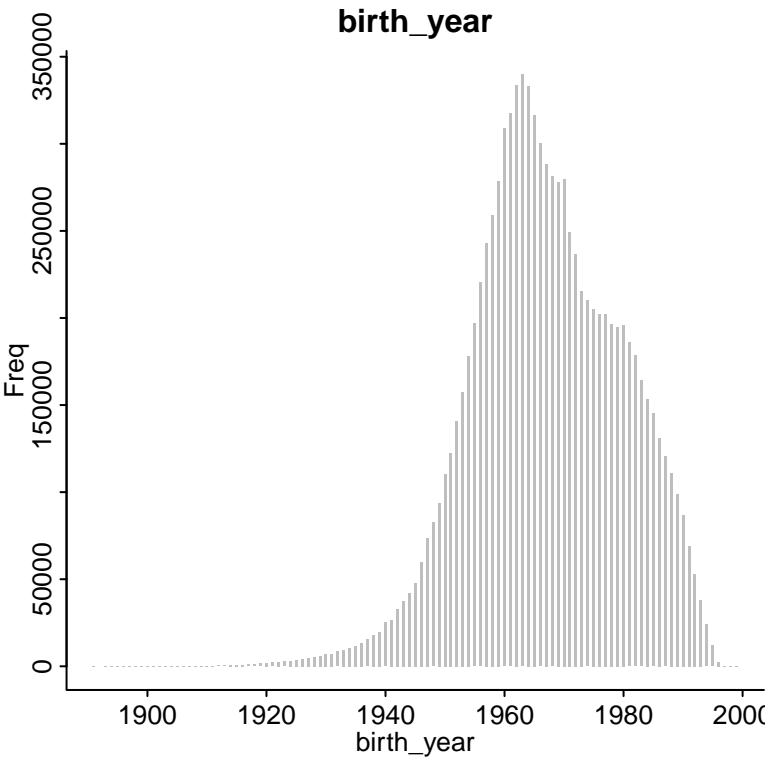
birth_year

Description

Individual's birth year.

Prompt

Enter the month/day/year (MM/DD/YYYY) of the subject's birth.

Summary

Name	Value
Min.	1891.000
1st Qu.	1959.000
Median	1966.000
Mean	1967.062
3rd Qu.	1976.000
Max.	1999.000

birth_year

birth_month

Description

Individual's birth month.

Prompt

Enter the month/day/year (MM/DD/YYYY) of the subject's birth.

Summary

Name	Value
Min.	1.000000
1st Qu.	4.000000
Median	7.000000
Mean	6.568342
3rd Qu.	10.000000
Max.	12.000000
NA's	66806.000000

birth_month
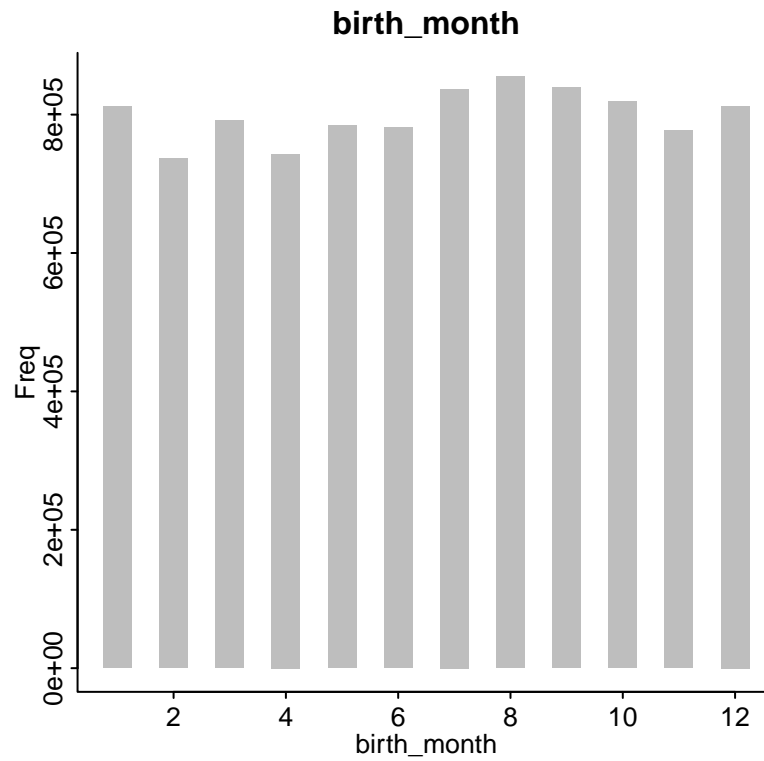
birth_day

Description

Individual's birth day.

Prompt

Enter the month/day/year (MM/DD/YYYY) of the subject's birth.

Summary

Name	Value
Min.	1.00000
1st Qu.	8.00000
Median	16.00000
Mean	15.62725
3rd Qu.	23.00000
Max.	31.00000
NA's	67009.00000

birth_day
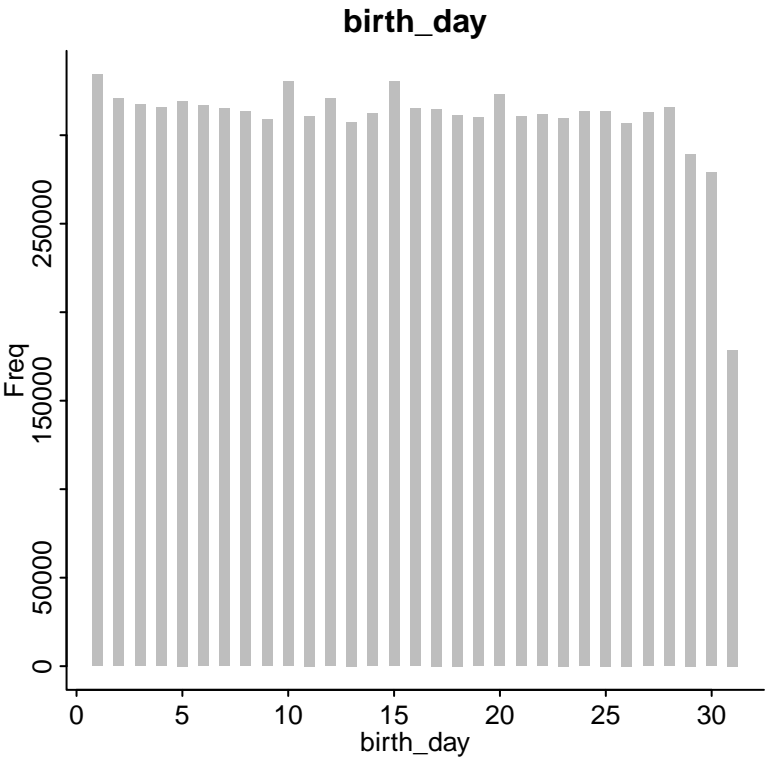
gender

Description

Individual's sex/gender.

Labels

- M = Male
- F = Female
- X = Unknown

Prompt

Enter the appropriate sex code abbreviation for the subject.

Summary

Name	Freq
male	7861805
female	1787784
unknown	5405

race

Description

Individual's race/ethnicity.

Labels

see race_or_ethnicity.csv

Prompt

Enter the appropriate race code abbreviation for the subject. The criteria for determining race categories are in the TABLES Section.

Notes

The codes for Asian/Pacific Islander became more detailed in 1991. San Francisco did not change its reporting practices until 2012, when it adopted the FBI's categories for race: white, black, American Indian, other Asian, and other. Since 2012, San Francisco has not distinguished between non-Hispanic whites and Hispanic whites.

Summary

Name	Freq
white	3504045
hispanic	3241557
black	2358065
other	164450
unknown	132547
american_indian	48894
other_asian	48371
filipino	47441
vietnamese	26384
chinese	26169
pacific_islander	19560
laotian	8834
japanese	5617
samoan	5562
asian_indian	5374
korean	5111
cambodian	2890
hawaiian	2789
guamanian	1334

deceased

Description

Whether individual is deceased, alive, or unknown.

Labels

- 0 - not deceased or no information
- 1 - deceased
- 2 - deceased (homicide)

Summary

Name	Freq
not_deceased_or_no_information	9591048
NA	55556
deceased_homicide	5756
deceased	2634

arrest_record_id

Description

where final disposition received or last reported step in criminal justice process.

Labels

- 10 - court disposition
- 11 - prosecutor
- 12 - police

Summary

Name	Freq
final_court_disposition	7426045
prosecutor	1701337
police	527612

arrest_bypass

Description

Internal processing code.

Summary

Name	Freq
no_bypass_needed	9618967
received_and_corrected	19947
NA	10975
not_corrected	5105

arrest_converted_data

Description

Internal processing code.

Summary

Name	Freq
not_converted	9515110
converted_1	139884

arrest__event__year

Description

Arrest year.

Prompt

Enter the month/day/year (MM-DD-YYYY) of: 1) the subject's arrest as it was submitted on the arrest fingerprint submission, 2) the date the complaint was sought or filed, or 3) the date of citation.

Summary

Name	Value
Min.	1929.000
1st Qu.	1991.000
Median	1998.000
Mean	1998.368
3rd Qu.	2006.000
Max.	2015.000

arrest_event_year
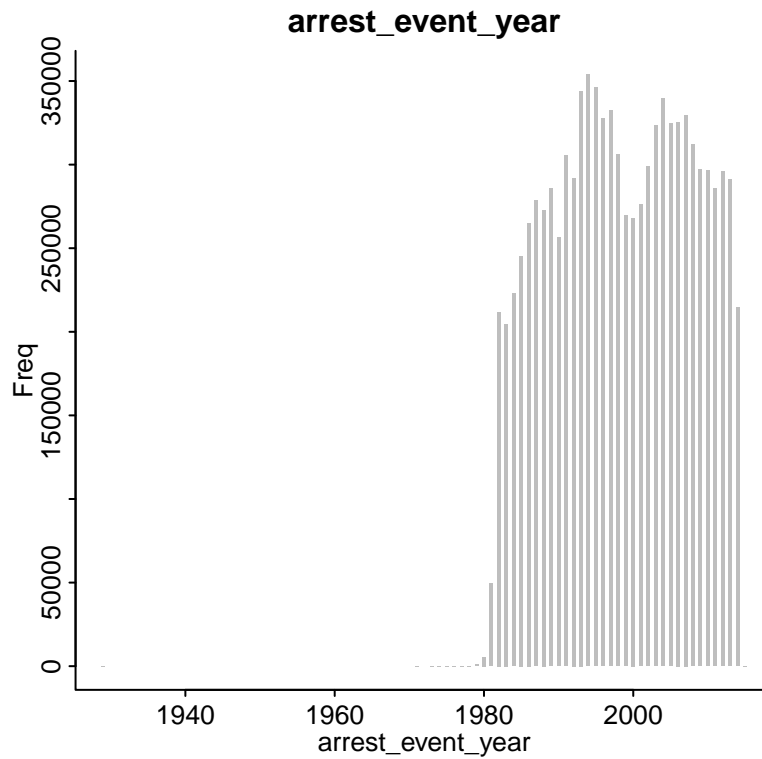
arrest__event__month

Description

Arrest month.

Prompt

Enter the month/day/year (MM-DD-YYYY) of: 1) the subject's arrest as it was submitted on the arrest fingerprint submission, 2) the date the complaint was sought or filed, or 3) the date of citation.

Summary

Name	Value
Min.	1.000000
1st Qu.	3.000000
Median	6.000000
Mean	6.411147
3rd Qu.	9.000000
Max.	12.000000

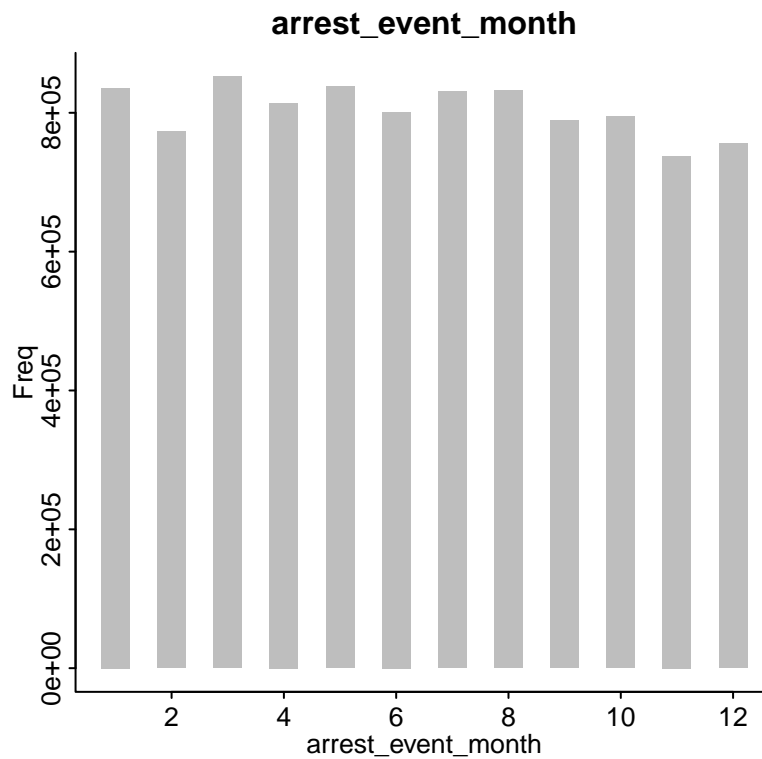arrest_event_month

arrest__event__day

Description
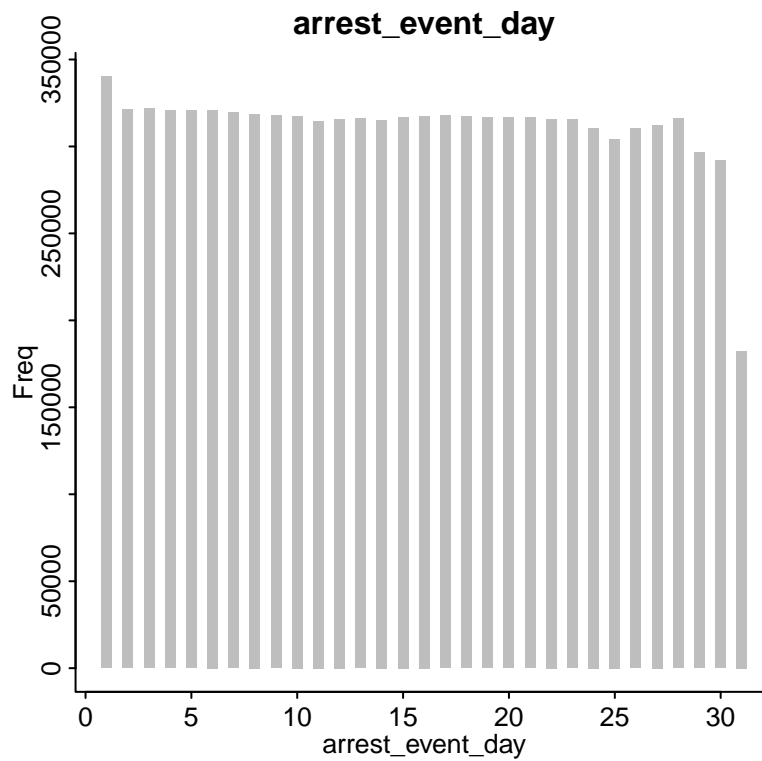
Arrest day.

Prompt

Enter the month/day/year (MM-DD-YYYY) of: 1) the subject's arrest as it was submitted on the arrest fingerprint submission, 2) the date the complaint was sought or filed, or 3) the date of citation.

Summary

Name	Value
Min.	1.0000
1st Qu.	8.0000
Median	16.0000
Mean	15.6296
3rd Qu.	23.0000
Max.	31.0000

arresting_agency

Description

Arresting or booking agency.

Labels

- PD Police Department
- SO Sheriff's Office (Department)
- CHP California Highway Patrol
- MR Marshal
- INS United States Immigration and Naturalization Service
- FBI Federal Bureau of Investigation
- DA District Attorney
- CA City Attorney
- NTF Narcotic Task Force

Prompt

Enter the complete name of the agency that arrested, cited, or requested a complaint be filed against the subject. Abbreviates for Type of Agency are acceptable; however, the county or city name should be spelled out followed by SO, PD, etc.

Notes

may actually be booking and not arresting agency

Summary

Name	Freq
194200	829442
380100	411893
190000	405271
360000	334991
371100	271722
431300	235863
010900	226795
330000	210015
300000	197450
340000	186042
340400	164186
370000	144609
100500	143750
150000	129940
150200	125882
194100	120693
390500	90896
300100	70622
500200	69642
390000	68235
010000	67389
490001	63597

Name	Freq
100000	58186
301900	58059
...	...

local_file_number

Description

Sealed.

arrest_charge_type

Description

Arrest offense level.

Labels

- F - felony
- M - misdemeanor
- I - infraction
- X - unknown

Summary

Name	Freq
felony	9653751
misdemeanor	1243

multiple_arrest_charges

Description

Number of arrest charges.

Labels

1-9 where 9 is 9+; prior to 1986 only used: 1 - one offense and 2 - more than one

Summary

Name	Freq
1	5711246
2	2269846
3	978276
4	390492
5	146942
6	61679
NA	55556
7	22838
8	10407
9	7712

arrest_disposition_year

Description

Date of final disposition decided by arresting officer.

Summary

Name	Value
Min.	1944.00
1st Qu.	1989.00
Median	1997.00
Mean	1997.35
3rd Qu.	2006.00
Max.	2015.00
NA's	7426566.00

arrest_disposition_year
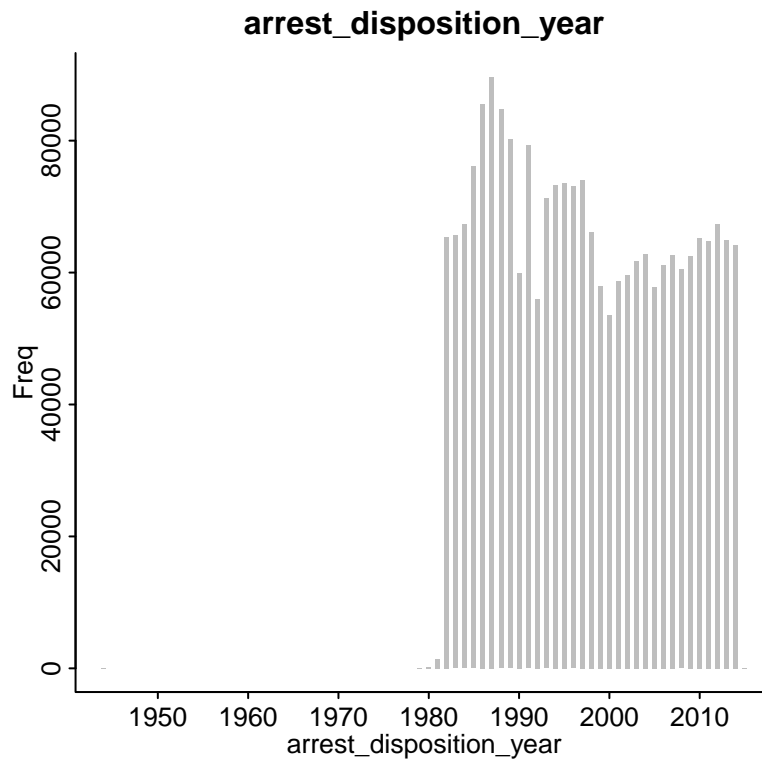
arrest__disposition__month

Description

Date of final disposition decided by arresting officer.

Summary

| Name	Value
Min.	1.000000e+00
1st Qu.	3.000000e+00
Median	6.000000e+00
Mean	6.345454e+00
3rd Qu.	9.000000e+00
Max.	1.200000e+01
NA's	7.426566e+06

arrest_disposition_month

27
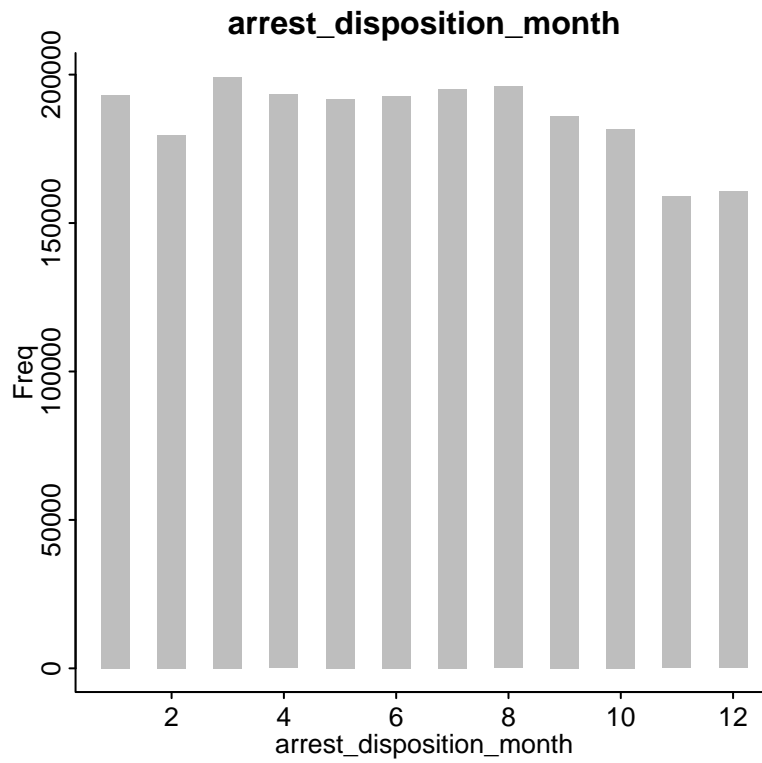
arrest_disposition_day

Description

Date of final disposition decided by arresting officer.

Summary

Name	Value
Min.	1.000000e+00
1st Qu.	8.000000e+00
Median	1.600000e+01
Mean	1.564913e+01
3rd Qu.	2.300000e+01
Max.	3.100000e+01
NA's	7.426566e+06

arrest_disposition_day
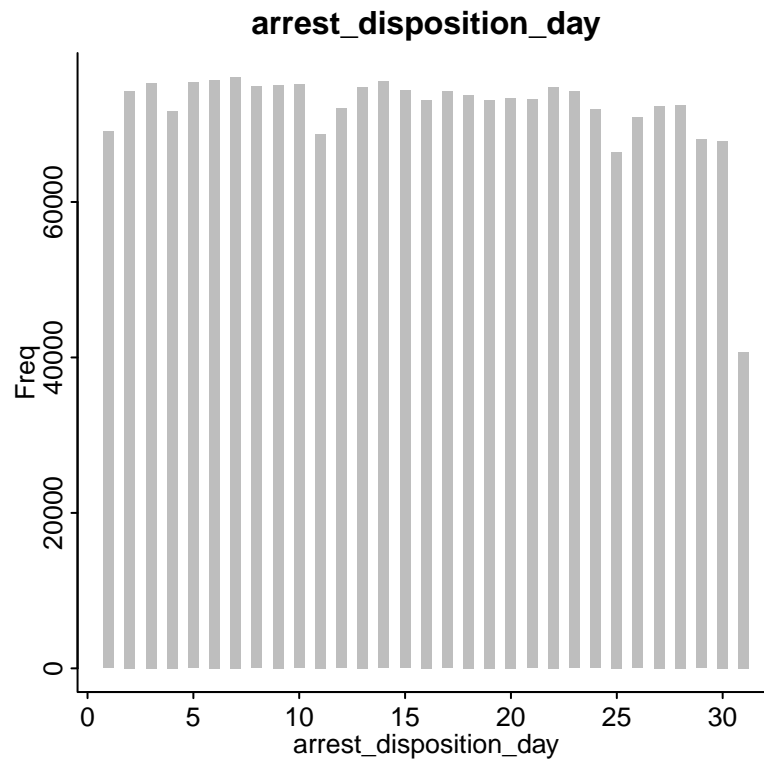
arrest_disposition_type

Description

Release or decline to prosecute rationale.

Labels

see dispositions.csv table

Notes

prosecutor rejects is only the final disposition if the Court Record ID is blank

Summary

Name	Freq
NA	7426045
da_dec_lack_of_cause	720714
da_dec_other	429302
lea_rel_other	196243
da_dec_lack_of_corpus	147449
da_dec_interest_of_justice	112572
lea_rel_lack_of_corpus	108734
da_dec_complaintant_refuses	100876
lea_rel_inadmissable_search	81336
da_dec_inadmissable_search	54007
lea_rel_further_investigation	52722
da_dec_probation_revoked_1203_2PC	50161
lea_rel_complaintant_refuses	47626
da_dec_prosecution_prefiling_deferred	37673
da_dec_deferred_recovation_of_parole	23611
lea_rel_unspecified	22246
lea_rel_arrestee_exonerated	18701
da_dec_combined_cases	13398
da_dec_witness_unavailable	11578

arrest_multiple_dispositions

Description

Number of arrest dispositions.

Labels

- 0 - subject did not receive all final dispositions at arrest level
- 1-9 where 9 is 9+;
- pre-1986 coding: 1 - one offense and 2 - more than one

Summary

Name	Freq
0	7549331
1	2016556
NA	55556
2	32456
3	1075
4	20

prior_record_code

Description

Number of prior prison commitments.

Labels

- 0 - no prior record
- 1 - misc prior record
- 7 - one prior prison commitment
- 8 - two prior prison commitments
- 9 - three or more prior prison commitments
- blank - pre-1982 means lengthy record

Summary

Name	Freq
misc_prior_record	4548209
length_record	2305492
no_prior_record	2087482
one_prior_commitment	425620
two_prior_commitments	164598
three_or_more_prior_commitments	123593

arrest_offense

Description

Offense for which individual was arrested.

Labels

see cjis_codeTable.csv

Summary

Name	Freq
22004	892330
35172	722847
35152	584825
38028	519683
38025	389090
13094	367757
12004	314317
24054	209856
23097	193583
35186	183162
35127	166715
35138	160973
28025	153981
35188	143985
16001	143112
35154	122654
35254	113757
35137	108170
52096	102668
24053	101858
28012	92609
35215	92047
35276	88215
28011	86967
. . .	. . .

arrest_qualifier

Description

Arrest offense qualifier.

Labels

see qualifier_codes.csv

arrest_summary_code

Description

BCS summary code for arrest offense.

Labels

see bcs_offense_codes_variables.csv

Summary

Name	Freq
6	2000302
12	1328695
14	1176712
8	1136148
9	1084953
10	419789
5	416760
25	408080
19	351713
13	345788
11	277054
20	205080
18	87613
16	81499
4	71178
15	57398
1	54755
7	50104
21	23614
24	21942
17	20654
22	19537
3	7487
23	6798
2	1341

arrest_edit_error_code

Description

Internal processing code.

court_record_id

Description

Court of final disposition.

Labels

- 20 - final disposition in superior court
- 21 - final disposition in lower court

Summary

Name	Freq
superior_court	4808342
lower_court	2965568
NA	1881084

court_bypass

Description

Internal processing code.

Summary

Name	Freq
no_bypass_needed	9184652
not_corrected	275274
NA	110311
corrections_sent	84757

court_converted_data

Description

Internal processing code.

Summary

Name	Freq
NA	9599438
1	55556

court__event__year

Description

Last date of court proceedings.

Notes

Seems to "root" the record as is the only year value that is the same as the year of the flat file from which the data are drawn.

Summary

Name	Value
Min.	1982.000
1st Qu.	1991.000
Median	1999.000
Mean	1998.767
3rd Qu.	2007.000
Max.	2014.000

39
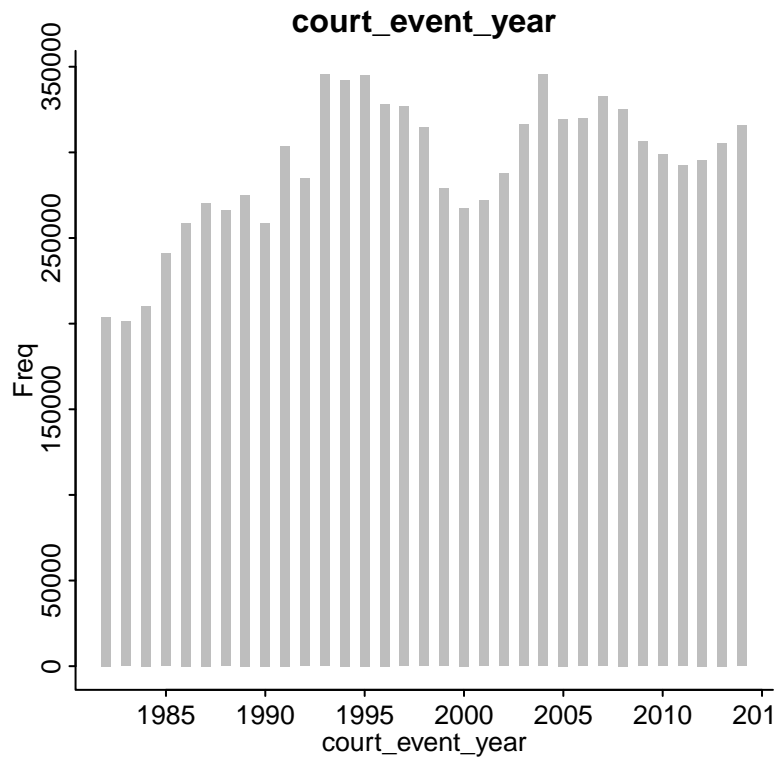
court_event_month

Description

Last date of court proceedings.

Summary

Name	Value
Min.	1.00000
1st Qu.	3.00000
Median	6.00000
Mean	6.31144
3rd Qu.	9.00000
Max.	12.00000

court_event_month
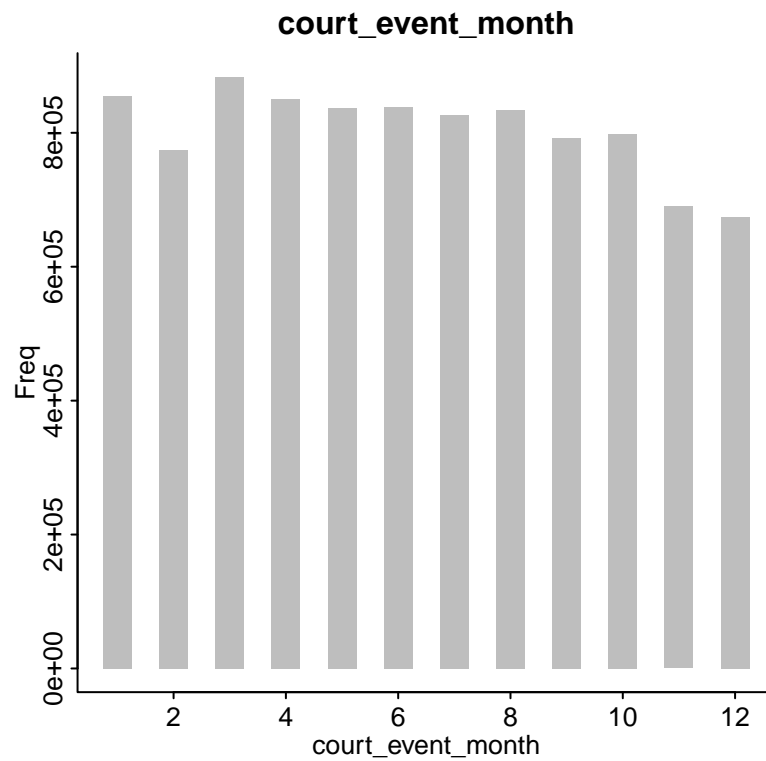
court_event_day

Description

Last date of court proceedings.

Summary

Name	Value
Min.	1.00000
1st Qu.	8.00000
Median	16.00000
Mean	15.62613
3rd Qu.	23.00000
Max.	31.00000

court_event_day

41
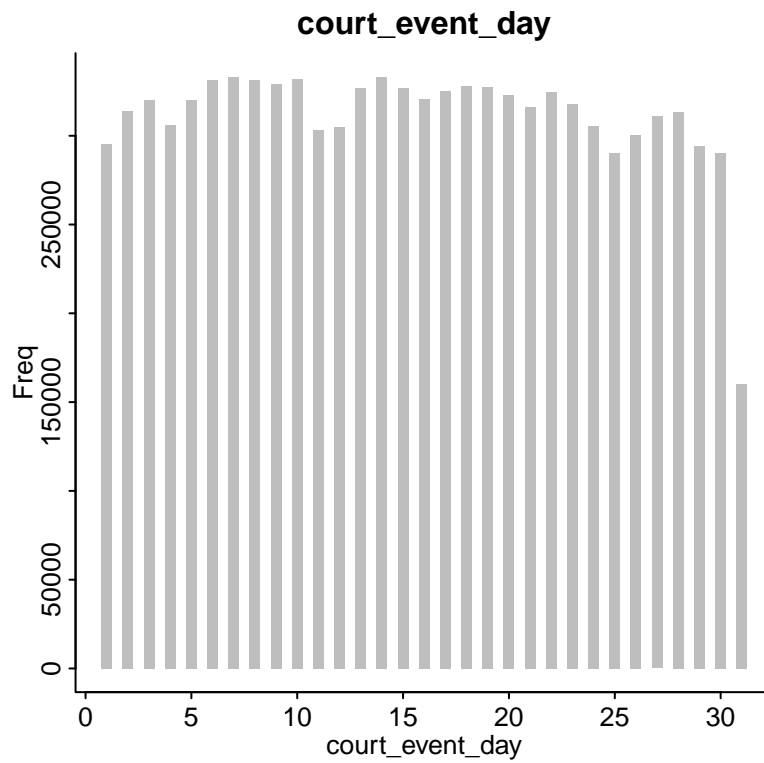
court_judicial_district

Description

Court judicial district.

Labels

see judicial_districts.csv

Summary

Name	Freq
19****	785494
19100J	304714
19463J	257809
37100J	211884
34470J	207425
43100J	190563
38****	177816
37****	167027
15420J	163929
10440J	150352
34100J	149181
38100J	144353
01100J	141132
01****	133451
43470J	129684
33460J	124430
30100J	119732
39460J	117005
30470J	115486
50100J	113418
30420J	111595
36470J	107659
49460J	101303
01440J	100452
. . .	. . .

court_number

Description

Sealed.

court_disposition_type

Description

Outcome of court trial.

Labels

- 21 - convicted
- 22 - acquitted
- 23 - not guilty, insane
- 24 - certified to juvenile court
- 25 - defendant deceased
- 26 - dismissed
- 27 - diversion dismissed (begins 1988)

Summary

Name	Freq
convicted	6416370
NA	1881084
dismissed	1172645
diversion_dismissed	155270
acquited	24680
certified_to_juvenile_court	2302
defendant_deceased	1581
not_guilty_insane	1062

court_multiple_dispositions

Description

Number of court dispositions.

Labels

Pre 1986:

- 1 - one disposition
- 2 - more than one

post 1985:

- 1 - only one
- 2 - 8 - two to eight
- 9 - 9+ dispositions

Summary

Name	Freq
1	3881016
2	3387667
NA	1921751
3	425894
4	35492
5	2866
6	278
7	29
9	1

court_proceeding_type

Description

Court proceeding type.

Summary

Name	Freq
misdemeanor_complaint_filed	2863867
felony_complaint_filed	2801467
NA	1887273
information	1136298
certification	948677
filed_as_infraction	10257
indictment	7155

sentence_type

Description

Court sentence type.

Summary

Name	Freq
felony_convicted_offense_felony_sentence	3325469
NA	3308608
misdemeanor_convicted_offense	1984192
filing_as_17b4_PC	661199
felony_convicted_offense_misdemeanor_per_17C	253836
felony_convicted_offense_misdemeanor_sentence	89892
infraction_convicted_offense	31798

sentence

Description

Court sentence.

Labels

- 01 - death
- 02 - prison
- 03 - CYA
- 04 - CRC
- 05 - MDSO
- 06 - probation and jail
- 07 - probation
- 08 - probation
- 09 - fine
- 10 - other: sentence suspended; work program; volunteer program; no sentence
- 11 - prison term suspended (post 1984)

Summary

Name	Freq
probation_and_jail	3889758
NA	3238624
prison	1175414
probation	819844
jail	367954
other	69925
fine	67473
CRC	20534
CYA	3710
death	883
suspended	726
MSDO	149

court_charge_type

Description

Offense level of conviction crime.

Labels

- F - felony
- M - misdemeanor
- I - infraction
- N - place holder codes
- X - unknown level of arrest

Summary

Name	Freq
felony	5249184
misdemeanor	2233766
NA	1876011
placeholder	285871
infraction	10162

court_multiple_disposition_charges

Description

Number of disposition charges.

Labels

Pre 1986:

- 1 - one offense
- 2 - more than one

post 1985:

- 1 - only one
- 2 - 8 - two to eight
- 9 - 9+ offenses

Notes

EG: is this always the number of conviction charges, or are there charges associated with non-conviction dispositions? Ask Alyson?

Summary

Name	Freq
1	3580934
2	2620343
NA	1921751
3	973389
4	358149
5	123652
6	45780
7	17987
8	7357
9	5652

court_disposition_offense

Description

Conviction offense.

Labels

see cjis_codeTable.csv

Notes

EG: todo

Summary

Name	Freq
NA	1890900
35172	544530
35152	417506
66085	285592
22004	247954
38028	195049
13094	170191
24054	149101
22022	138002
13045	137124
23097	129421
35138	120468
38064	116455
35353	97894
28025	96752
35154	96101
35127	91976
12004	90088
35186	85894
22021	77740
35212	76463
23132	72771
23099	71337
35276	68789
. . .	. . .

court_qualifier

Description

Conviction offense qualifiers.

Labels

see qualifier_codes.csv

Notes

EG: todo

Summary

Name	Freq
	9340440
AB	71533
KA	38453
XX	35277
MJ	30472
ML	17468
MC	17293
ME	13615
MB	10812
MD	9223
MM	9222
MW	7034
MK	6437
NS	6131
ON	2266
MLMJ	1888
MJML	1812
MT	1766
XXMC	1453
MF	1163
AE	1117
ABME	997
MA	936
MU	887
. . .	. . .

court_summary_code

Description

Conviction offense summary code.

Labels

see bcs_offense_codes_variables.csv

Summary

Name	Freq
NA	1890901
12	838025
14	831928
9	783351
6	722912
30	550353
8	541489
60	431875
31	310474
36	291119
25	226112
10	218266
19	207727
13	189618
11	180195
5	177660
51	130717
32	90625
50	89011
20	82303
46	80839
47	65756
34	54448
16	53885
. . .	. . .

court_edit_error_code

**Description**

Internal processing code.

Data Cleaning

Court Event Date

Most of the single-digit births are from prior to 1992, when years were stored as only 2 digits. For reasons unknown, 1995 is also in a 2- digit year format. The court event year match that of the corresponding file, so that they range from 82 to 2014, and the missing digits can be safely added. 4 are unable to be parsed into a correct date:

court_event_year	court_event_month	court_event_day	court_event_date
1982	2	29	NA
1982	2	29	NA
1982	11	31	NA
1982	11	31	NA

Arrest Event Date

Arrest event codings appear to be more inconsistent. For all but 1982, arrest years go back no further than 1979. 1 record lie outside of this range. Nevertheless, it appears as if the two-year coded years can safely be converted to four-year codes. After doing so, 1 record fail to parse.
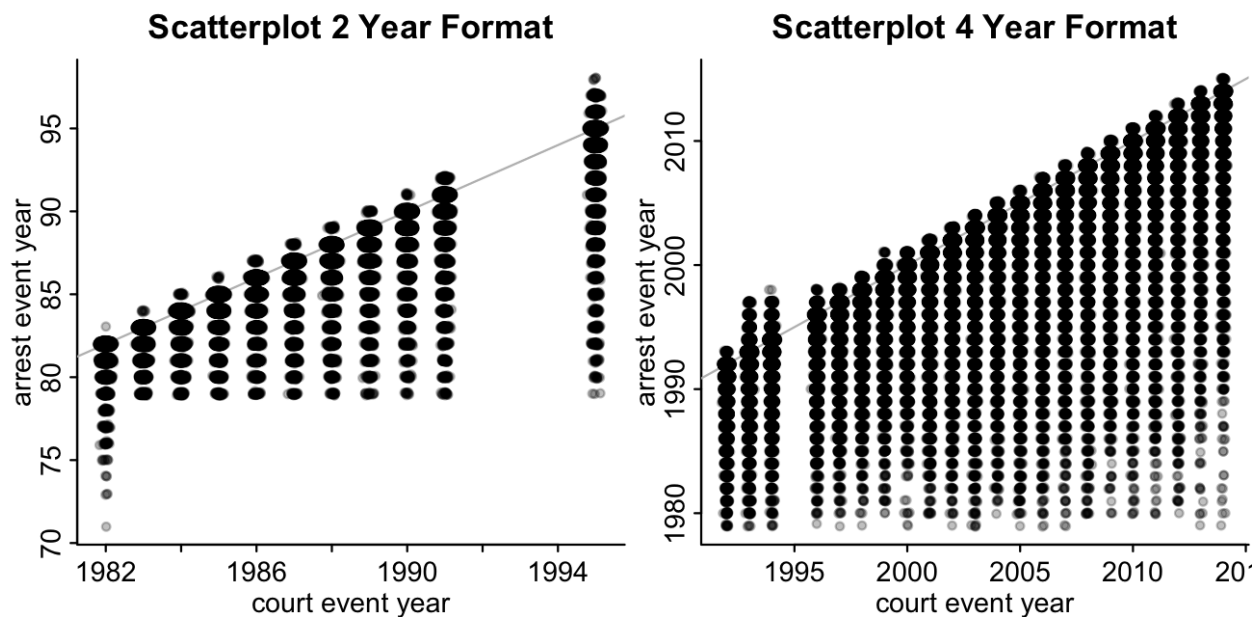
court_event_date	arrest_event_year	arrest_event_month	arrest_event_day	arrest_event_date
1982-11-05	1981	2	29	NA

Birth Date

The coding for birth years has numerous anomalies. 1982 has an abnormally large number of birth years equal to 99, while other years have clusters at 0.

Scatterplot 2 Year Format

birth year vs court event year

Scatterplot 4 Year Format

birth year vs court event year
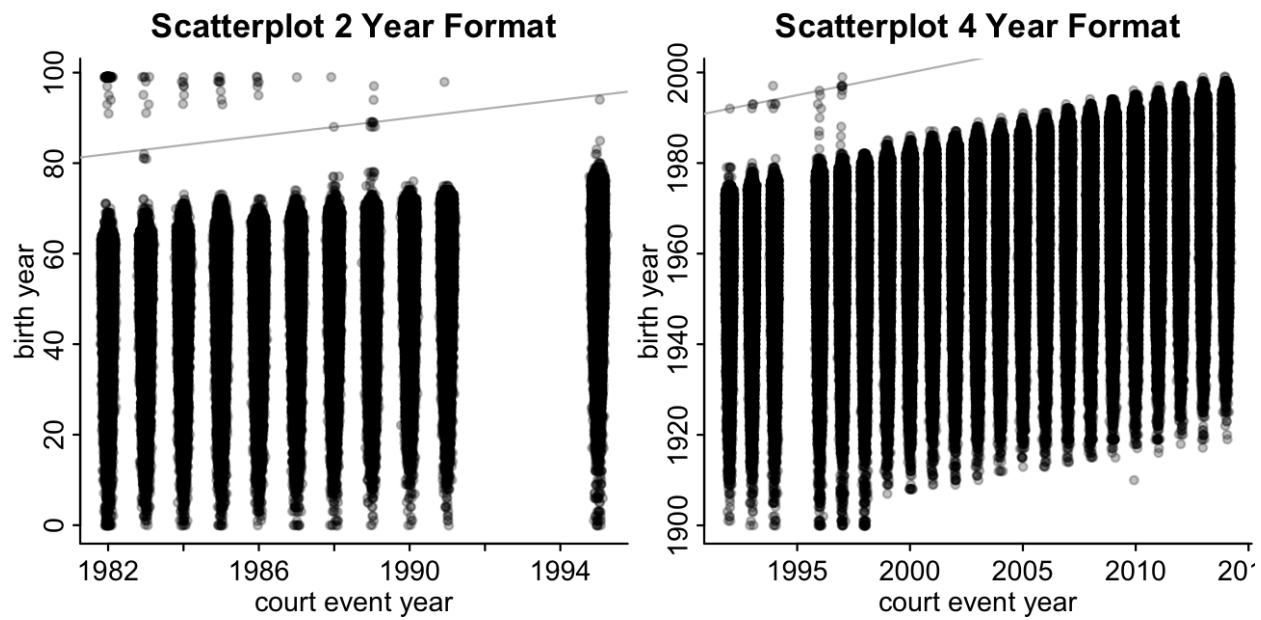
birth_year	court_event_year 1982	1983	1984	1985	1986	1987	1988
4	14	10	7	13	4	9	1
3	13	5	6	4	8	5	4
2	11	6	4	5	2	0	0
1	16	3	7	2	2	1	4
0	19	2	5	8	1	7	5
99	93	4	1	2	2	1	1
98	0	1	2	4	2	0	0
97	1	0	2	0	1	0	0
96	0	0	0	1	0	0	0
95	1	1	1	0	1	0	0

In addition, there are a number of records where the birth-year is extraordinarily close to the court-event-year, including in the near-future.

birth_year	court_event_year 1983	1988	1989	1992	1993	1994	1995	1996	1997
81	3	0	0	0	0	0	0	0	0
82	1	0	0	0	0	0	0	0	0
88	0	1	3	0	0	0	0	0	0
89	0	0	5	0	0	0	0	0	0
94	0	0	0	0	0	0	1	0	0
1992	0	0	0	1	1	1	0	0	0
1993	0	0	0	0	2	3	0	0	0
1995	0	0	0	0	0	0	0	1	1
1996	0	0	0	0	0	0	0	1	1
1997	0	0	0	0	0	0	0	0	4
1999	0	0	0	0	0	0	0	0	1

To impute the century for the birth-year, we follow the rule:

1. If the listed 2-digit birth year is greater than 3 years from the court-event-year minus 1900, then we assume that it is in the 1800s.
2. Otherwise, we assume that it is in the 1900s.

This scheme allows the birth years that are close to the court years to persist, as that trend exists in the data sets for 4-digit years. These records bear further investigation as they seem unlikely to be correct.

After making this correction, a number of birth dates still cannot be successfully parsed, largely owing to a number of records that have either the birth month or birth day recorded as 0 (66931 rows). The remaining rows all seem to be coded as 99/99/99 and are in the court year 1982, indicating a particular kind of not-available.

court_event_date	birth_year	birth_month	birth_day
1982-02-24	1899	99	99
1982-01-28	1899	99	99
1982-12-02	1899	99	99
1982-12-02	1899	99	99
. . .	. . .	. . .	. . .

We can look at the distribution of ages in the data by subtracting the court event date from the birth date.
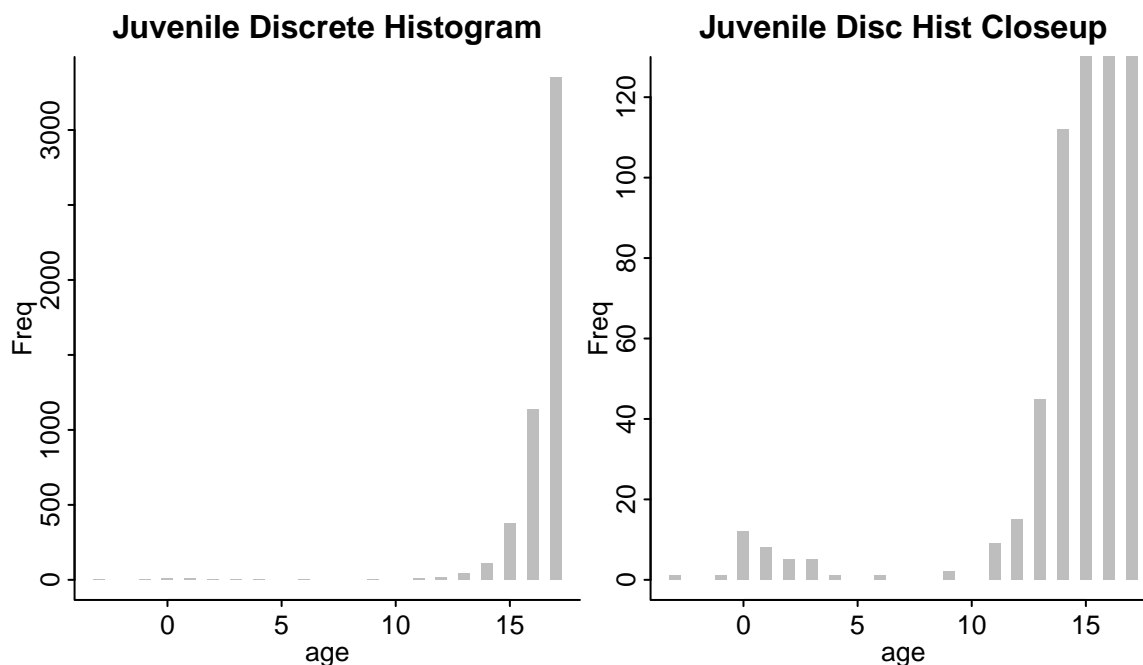
Arrest Disposition Date

The disposition dates can all safely be assumed to be in the same century as the court event, as it seems rare that the disposition date happens significantly after. Excluding those that are naturally missing, 2 dates cannot be parsed.

Scatterplot 2 Year Format

Scatterplot 4 Year Format
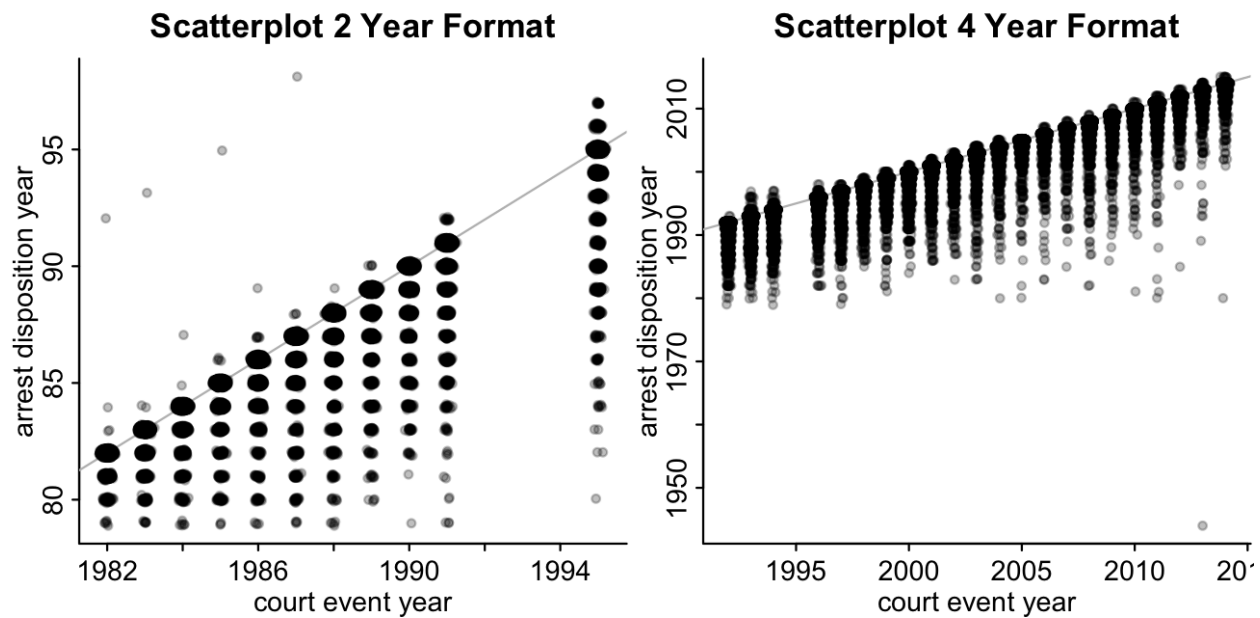
crt_evt_yr	crt_evt_mo	crt_evt_dy	arr_dsp_yr	arr_dsp_mo	arr_dsp_dy
1982	2	29	1982	2	29
1982	2	29	1982	2	29

Analyses

Table Joins

Some entries in the OBTS do not have matching values in the related tables, specifically the judicial districts, arresting agencies, and CJIS codes.

Judicial Districts

A total of 575 rows have listed judicial districts that cannot be matched.

Unmatched by Jurisdiction Tabulation

Name	Freq
29640J	369
50675J	74
04650J	54
31655J	19
56450J	10
33640J	5
37465J	5
08650J	4
28650J	3
32620J	3

Name	Freq
07110J	2
27660J	2
48680J	2
07645J	1
10671J	1
13640J	1
15610J	1
15631J	1
15640J	1
15660J	1
15675J	1
24680J	1
27680J	1
29440J	1
. . .	. . .

Unmatched by Court Event Year Tabulation

Name	Freq
1982	35
1983	25
1984	5
1985	7
1986	44
1987	62
1988	115
1989	174
1990	67
1991	8
1992	3
1993	3
1994	1
1995	4
1996	2
1997	2
1999	2
2000	1
2001	1
2002	1
2003	1
2004	3
2007	1
2008	1
2009	1
2011	4
2012	2

Arresting Agencies

A total of 66869 rows have listed arresting agencies that cannot be matched.

Unmatched by Agency Tabulation

Name	Freq
191500	13649
379900	11124
999999	6863
340001	2961
159900	2716
339900	2413
340003	1904
449900	1741
210400	1672
399900	1437
509DEU	1268
430300	1142
339931	916
109900	891
419900	793
569966	781
509900	651
169NTF	634
129900	633
450NTF	599
309900	529
019900	510
210800	462
209900	455
. . .	. . .

Unmatched by Court Event Year Tabulation

Name	Freq
1982	1518
1983	2110
1984	2747
1985	3854
1986	4067
1987	3198
1988	2482
1989	2097
1990	2421
1991	2576
1992	2077
1993	4112
1994	3634
1995	3542
1996	2911

Name	Freq
1997	2891
1998	2789
1999	2641
2000	1929
2001	1040
2002	1534
2003	1849
2004	1781
2005	1858
2006	1192
2007	373
2008	483
2009	492
2010	563
2011	490
2012	628
2013	455
2014	535

Offense Codes

A total of 62159 rows have arrest offense codes that cannot be matched, while 358603 rows have court disposition offense codes that cannot be matched.

Unmatched by Offense Code Tabulation

code	arrest_offense	court_disposition_offense
66085	0	285592
13998	48776	11229
66038	0	22189
54054	0	9517
66001	0	8479
13187	0	4966
66065	209	4713
42000	3316	262
66066	0	3346
52041	1936	913
42999	0	1569
66000	1178	355
24052	1406	112
35999	0	928
52000	836	66
58007	706	164
62000	468	309
24008	564	3
13042	374	173
66052	12	531
53085	440	52
13997	10	371
66999	0	351

code	arrest_offense	court_disposition_offense
50052	0	295
...	...	...

Unmatched by Court Event Year Tabulation

court_event_year	arrest_offense	court_disposition_offense
1982	5410	4239
1983	1904	4257
1984	1925	4393
1985	2004	4770
1986	1884	4469
1987	1581	4281
1988	1973	1647
1989	2348	1943
1990	3532	2139
1991	3122	2485
1992	2857	2119
1993	2662	13413
1994	2052	10160
1995	2188	11202
1996	2115	9549
1997	1770	22302
1998	1842	10385
1999	1752	8783
2000	1591	13007
2001	1309	17955
2002	1358	39712
2003	1187	10174
2004	1432	13767
2005	1228	11130
2006	1347	16083
2007	1088	14488
2008	1186	14013
2009	1169	14152
2010	1124	12492
2011	1138	12449
2012	1594	14537
2013	1266	15901
2014	1221	16207

Qualifier Codes

A total of 106154 rows have arrest qualifier codes that cannot be matched, while 229180 rows have court disposition offense codes that cannot be matched. Because rows can have multiple qualifier codes, the following tables do not necessarily add up to these numbers.

Unmatched by Offense Code Tabulation

code	arrest_qualifier	court_qualifier
AW	74318	140
XX	28688	37769
KA	13	39898
MJ	63	38754
ML	99	24701
MC	127	21172
ME	113	19466
MB	261	15846
MD	128	12675
MW	230	9474
MK	43	8350
NS	48	8111
MU	0	2307
AE	1091	1215
MT	9	2209
MF	13	1409
MS	0	1296
MA	32	1091
AR	944	7
MV	13	731
OA	11	691
NM	2	456
AA	49	396
AL	433	3
. . .	. . .	. . .

Unmatched by Court Event Year Tabulation

court_event_year	arrest_qualifier	court_qualifier
1982	1820	2806
1983	3474	4159
1984	4617	4736
1985	4252	5235
1986	3988	4500
1987	5426	4395
1988	5078	4656
1989	3777	4960
1990	3842	7421
1991	2787	7928
1992	3362	9574
1993	4293	9949
1994	4402	8990
1995	3345	9012
1996	2214	8129
1997	3526	6340
1998	4198	6587
1999	4524	5682
2000	3742	5181
2001	2783	4810
2002	2529	4533
2003	2812	4769

| court_event_year	arrest_qualifier	court_qualifier
2004	2634	5596
2005	2268	5500
2006	2337	5197
2007	1837	4753
2008	2545	4901
2009	2830	4668
2010	2754	5026
2011	2592	7095
2012	2008	16337
2013	1731	18466
2014	1827	17289

Arrest and Court Disposition Types

There are two disposition fields in the OBTS, one corresponding to the arrest segment and one corresponding to the court segment. In general, when an arrest disposition is present that implies that the subject was released and no court disposition should be present. Conversely, if a court disposition is present, the subject was *not* released and thus should have no arrest disposition. In fact, whenever one is missing the other is present, however 347865 records have *both*.

arrest_disp	court_disp acqui...	certi...	convi...	defen...	dismi...	diver...	not_g...	NA
da_dec_combi..	7	2	1649	0	334	18	0	11388
da_dec_compl..	5	1	1191	1	421	42	0	99215
da_dec_defer...	1	0	491	0	181	4	0	22934
da_dec_inadm..	0	0	801	1	223	36	0	52946
da_dec_inter...	29	1	12599	2	2509	263	1	97168
da_dec_lack_...	98	12	31949	17	8612	1033	1	678992
da_dec_lack_...	78	10	24211	4	6697	859	2	115588
da_dec_other	478	87	148218	46	32497	5921	11	242044
da_dec_proba..	0	0	907	1	590	4	0	48659
da_dec_prose..	0	0	349	0	144	17	0	37163
da_dec_witne..	0	0	300	0	90	4	0	11184
lea_rel_arre...	2	0	206	0	58	4	0	18431
lea_rel_comp...	11	1	2343	0	677	15	1	44578
lea_rel_furt...	26	0	5216	10	937	201	3	46329
lea_rel_inad...	5	0	1823	0	401	23	0	79084
lea_rel_lack...	6	1	2681	0	858	45	0	105143
lea_rel_other	63	6	28849	11	4845	1277	7	161185
lea_rel_unsp...	35	17	11732	0	1363	44	2	9053
NA	23836	2164	6140855	1488	1111208	145460	1034	0

Qualifier Codes Over Time

This section shows the usage of qualifier codes over time. The vast majority of codes are used < 10 times overall, while a few have usages in the thousands. The following histograms show the frequencies of total occurences, that is the codes are tabulated across all records and what is displayed is how many have those counts.
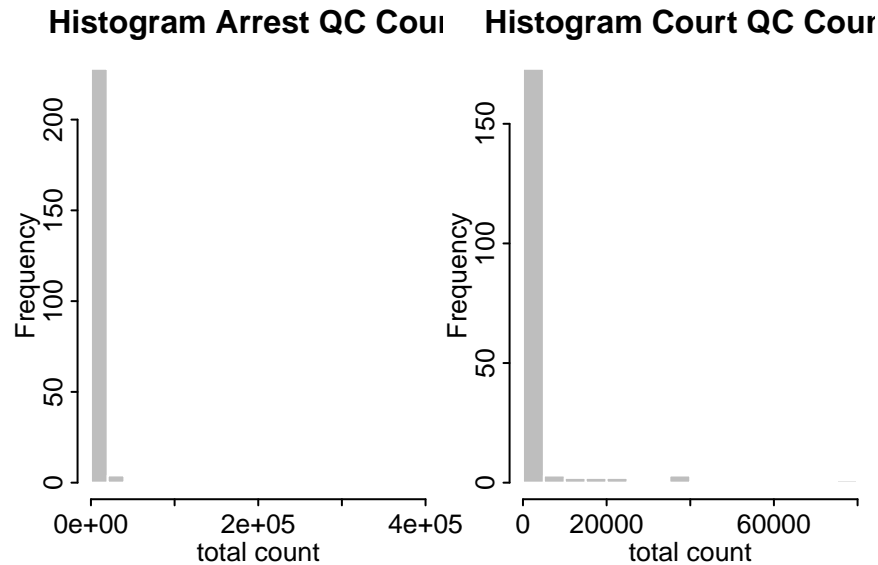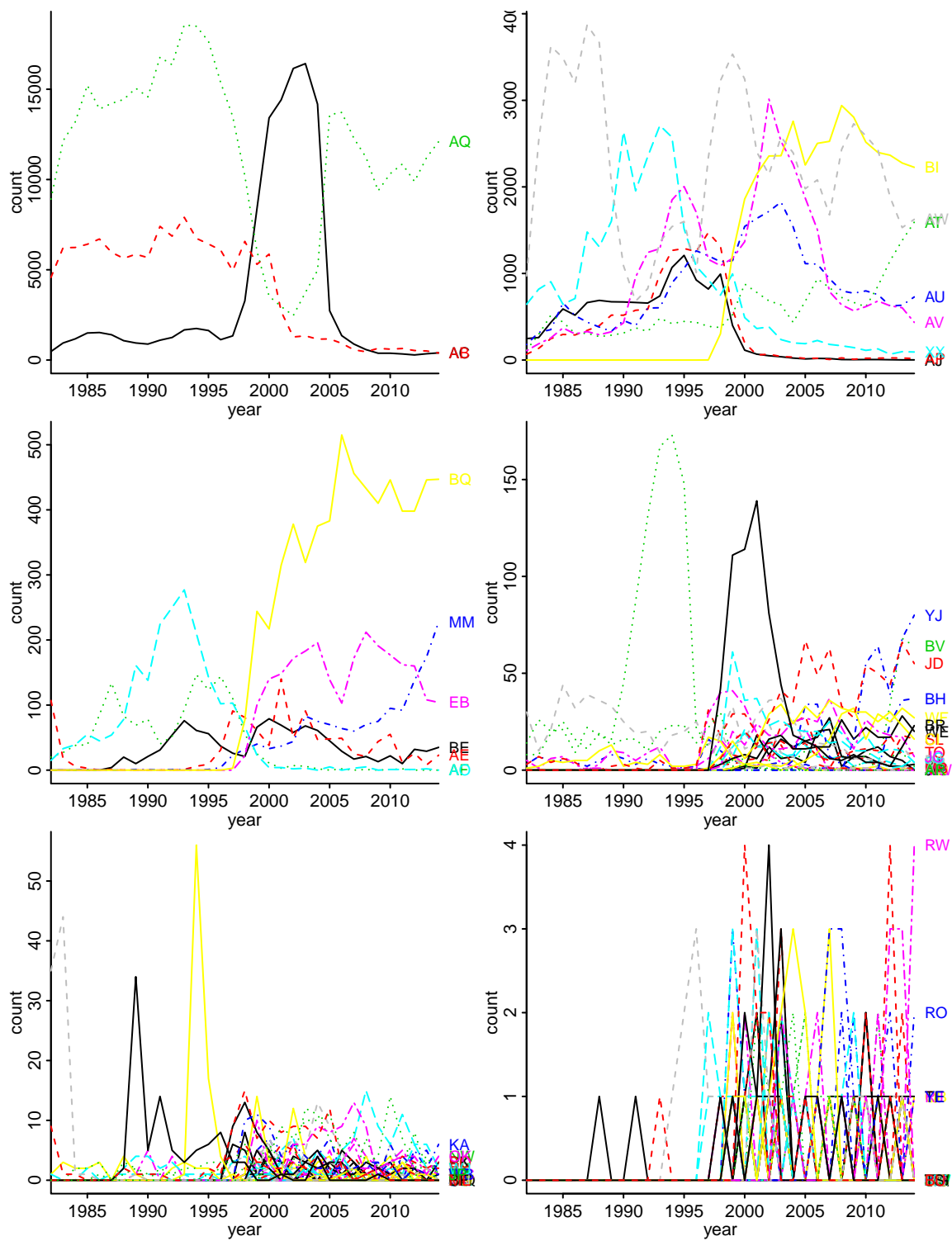

Codes with 500 or more usages in the arrest or court section are show in the table below. When a value is given as `NA`, it does not exist in the qualifier codes section of the OBTS Tech Manual.

qualifier	arrest_count	court_count	literal_short	type	literal
AQ	386809	0	ZZZZ	XX	WARRANT
AB	130297	78196	ZZZZ	XX	ATTEMPTED
AC	113670	0	ZZZZ	XX	BENCH WARRANT
AW	74318	0	NA	NA	NA
XX	28688	37769	NA	NA	NA
KA	0	39898	NA	NA	NA
MJ	0	38754	NA	NA	NA
BI	37824	0	ATTEMPT	PC	ATT TO COMMIT CRIME
AV	34626	0	ZZZZ	XX	COURT REMAND
AU	28470	0	ZZZZ	XX	BOOK / RELEASE
ML	0	24701	NA	NA	NA
MC	0	21172	NA	NA	NA
AT	19482	0	ZZZZ	XX	CRT ORDER BOOK
ME	0	19466	NA	NA	NA
MB	0	15846	NA	NA	NA
AP	12735	0	ZZZZ	XX	PROBATION VIOL
MD	0	12675	NA	NA	NA
AJ	12598	0	ZZZZ	XX	FAIL TO APPEAR
MM	1443	10732	PRIOR	PC	FEL/WHILE ON BAIL OR RECOG
MW	0	9474	NA	NA	NA
MK	0	8350	NA	NA	NA
NS	0	8111	NA	NA	NA
BQ	6260	0	PROBATION	PC	PROB VIOL/REARRESTED

qualifier	arrest_count	court_count	literal_short	type	literal
ON	0	3011	ZZZZ	XX	INVOLVE FIREARM
...	...	...	...	...	...

Finally, we show the usage of codes over time.

Arrest Qualifiers

Court Qualifiers