

Data file description

Contains data from 200812_pubnbr_app_code_pub_date_app_cou.dta

obs: 955,849
vars: 5 29 Jul 2009 19:29
size: 21,028,678 (96.0% of memory free)

variable name	storage type	display format	value label	variable label
app_code	long	%10.0g		App_code
app_cou	str2	%9s		App_cou
pubnbr	double	%10.0g		EPO patent publication number
yr_fil	int	%9.0g		filing year
yr_pr	int	%9.0g		priority year

Sorted by: pubnbr

Contains data from 200812_EPO_App_header.dta

obs: 199,291
vars: 14 29 Jul 2009 19:31
size: 85,495,839 (83.7% of memory free)

```
-----  
          storage  display  value  
variable name  type    format  label    variable label  
-----  
app_code       long    %10.0g  7 digit applicant code  
app_code_short float    %9.0g  6 digit applicant code  
app_name       str147  %147s   Original Applicant Name by the EPO  
std_name       str141  %141s   Harmonized Applicant Name by GTh's software: first 100 chars  
ctry_code      str2    %9s     Applicant's country code  
address_count  byte    %9.0g   # Addresses available  
town           str98   %98s    Applicant's town code  
notown         float   %9.0g   No town information available  
zip_code       str7    %9s     Applicant's zip code  
nozip          byte    %9.0g   No zip code available  
region         str6    %9s     Applicant's region code  
ind            float   %9.0g   Applicant being an Individual: 1=Individual 0=otherwise  
nbo            float   %9.0g   Applicant being an Non Business Org: 1=NBO 0=otherwise  
ep_npat        int     %12.0g  Frequency  
-----
```

Sorted by: app_code

Contains data from 200812_EPO_BVD_matching_090407.dta

obs: 70,668
vars: 8 29 Jul 2009 19:31
size: 18,939,024 (96.4% of memory free)

variable name	storage type	display format	value label	variable label
app_code	long	%10.0g		7 digit applicant code
bvdid	str19	%19s		BVD company ID of the subsidiary
country	str2	%9s		Company country
std_name	str135	%135s		Harmonized Applicant Name by GTh's software: first 100 chars
company	str100	%100s		Amadeus company name
match_loc	byte	%9.0g		Does location match? 1=yes; 0=no; -1=unknown
scorematch	byte	%14.0g	score_codes	A goodness score from 0,1,2,3,4,5,6,7,8,9
fuzzymatch	int	%9.0g		% of the number of tokens in common

Sorted by: app_code

Contains data from US_bvd_matching.dta

obs: 54,014
vars: 11 14 Mar 2010 12:35
size: 12,423,220 (97.6% of memory free)

variable name	storage type	display format	value label	variable label
uspto_code	long	%12.0g		6 digit assignee code
country	str2	%9s		Assignee country
std_name	str44	%44s		Harmonized Assignee Name by GTh's software: first 100 chars
us_npat	int	%8.0g		US granted patents in the USPTO
bvdid	str19	%19s		BVD company ID of the subsidiary
scorematch	byte	%14.0g	score_codes	A goodness score from 0,1-10
city_comp	str43	%43s		Company's city
match_loc	byte	%9.0g		Does location match? 1=yes; 0=no; -1=unknown
fuzzymatch	int	%9.0g		% of the number of tokens in common
company	str99	%99s		Amadeus company name
zip_comp	str9	%9s		Company's zip

Sorted by: uspto_code

Contains data from 200703_pubnbr_uspto_code_dates.dta

obs: 1,209,941
vars: 4 14 Mar 2010 15:15
size: 29,038,584 (97.2% of memory free)

	storage	display	value	
variable name	type	format	label	variable label
pubnbr	double	%12.0g		Publication number
uspto_code	long	%12.0g		6 digit assignee code
yr_fil	float	%9.0g		filing year
yr_pr	float	%9.0g		priority year

Sorted by: pubnbr uspto_code

NOTES: Missing dates for 227 patent documents.

Contains data from US_tm_main_file.dta

obs: 4,726,102

vars: 26 3 Apr 2009 03:13

size: 307,196,630 (41.4% of memory free)

storage	display	value		
variable name	type	format	label	variable label
serial_nbr	long	%12.0g		ID assigned to applications with a filing date
reg_nbr	long	%12.0g		ID assigned to applications with received approval for registration
filing_date	long	%12.0g		Date when a complete application was received by the USPTO
oppo_date	long	%12.0g		Date when the application was published for opposition
prio_date	long	%12.0g		Date of filing in a foreign country which becomes the priority filing date in US
reg_date	long	%12.0g		Date on which a mark was registered by the USPTO
renewal_date	long	%12.0g		Last date when a renewal for a mark was registered by the USPTO
aband_date	long	%12.0g		Date when the mark was abandoned. It applies only to dead marks
canc_date	long	%12.0g		Marks that were cancelled on a particular date
amended_reg	long	%12.0g		Amended_register
renewal_gen	byte	%8.0g		Renewal generation
date	float	%9.0g		Reference date for stats analysis, earliest date possible
year	int	%9.0g		Reference year for stats analysis, earliest year possible
abandoned	byte	%8.0g		1=abandoned; 0=otherwise
cancelled	byte	%8.0g		1=cancelled; 0=otherwise
expired	byte	%8.0g		1=cancelled; 0=otherwise
trademark	byte	%8.0g		Mark classified in the NICE classes 1-34

servicemark	byte	%8.0g	Mark classified in the NICE classes 35-45
dead	byte	%8.0g	1=dead; 0=otherwise
pending	byte	%8.0g	1=pending; 0=otherwise
registered	byte	%8.0g	1=registered; 0=otherwise
affids_10yr	byte	%8.0g	10th year affidavits filed
affids_6yr	byte	%8.0g	6th year affidavits filed
reg_changed	byte	%9.0g	Change in registration has occurred
int_class_total	byte	%8.0g	# of International Classification of Goods and Services (Nice Agreement)
us_class_total	int	%8.0g	# of US class information

Sorted by: serial_nbr

.