

TED CSV open data

NOTES & CODEBOOK

Version 3.1, 2018-02-23

Thank you for your interest in [Tenders Electronic Daily \(TED\) comma separated value \(CSV\) open data](#)¹. Before starting to work with the data, please read the notes and codebook below. They are necessary for drawing correct conclusions from the data.

For expert users, advanced notes on the dataset are also available on the open data website.

Table of Contents

1. COMMUNICATION	2
2. RELIABILITY & COVERAGE	2
3. STRUCTURE & INTERPRETATION.....	3
4. CODEBOOK.....	7
5. ANNEX I – CALCULATION OF VALUE_EURO_FIN_1	22
6. ANNEX II – CALCULATION OF DURATION.....	24
7. ANNEX III – VERSION HISTORY	25
8. ANNEX IV – DOWNLOAD SIZES.....	28

¹ TED data in XML is available from the [Publications Office of the European Union](#) at <https://open-data.europa.eu/en/data/dataset/ted-1>. All disclaimers which relate to TED XML data also apply to the TED comma separate value data and vice versa.

1. COMMUNICATION

1.1. The Commission is interested in the results of research on public procurement. We would be grateful to receive any output based on the data (e.g. papers, reports, links to applications) at GROW-G4@ec.europa.eu.

1.2. We recommend citing the dataset in the following format:

TED csv dataset (YYYY-YYYY), Tenders Electronic Daily, supplement to the Official Journal of the European Union. DG Internal Market, Industry, Entrepreneurship, and SMEs, European Commission, Brussels. Available at <https://open-data.europa.eu/cs/data/dataset/ted-csv>. Version 3.1. Accessed on YYYY-MM-DD.

1.3. To support the exchange of ideas, especially between practitioners and academics, the Commission hosts an [open wiki](#) with ideas for research questions and existing examples of reuse related to this dataset.

1.4. While spreadsheet programs (such as Excel) may be able to open a single year of data, they are not powerful enough to manage datasets with multiple years. We recommend using [specialized statistical tools](#), many of which are also available for free (e.g. [R](#) with [RStudio](#); [statistical packages](#) for [python](#) with [Jupyter](#)), including extensive online training courses and materials.

1.5. If you are still experiencing performance problems, dropping textual variables² – which are not used in most types of analyses – can be a useful first step.

2. RELIABILITY & COVERAGE

2.1. The data is provided "as is". The source of the data is unverified output from contracting authorities or entities across Europe. It is not uncommon for data to be input incorrectly (for examples see presentations available on the open data website) or be missing, and thus great care must be taken with data management and interpretation. Please note that due to resource constraints the European Commission is regrettably not able to provide support in analysing the data.

However, we are currently taking steps to clean historical data (mainly names of entities and values) to broaden opportunities for analysis, using both advanced techniques such as machine learning and manual cleaning. **If you are considering investing significant resources into data cleaning for your project, please contact us at the email address above, as we might be able to help each other out.**

2.2. The data comes from the European Economic Area, Switzerland, the Former Yugoslav Republic of Macedonia³, and EU institutions⁴ and covers the time

² E.g. TED_NOTICE_URL, CAE_NAME, CAE_NATIONALID, CAE_ADDRESS, CAE_TOWN, CAE_POSTAL_CODE, CRIT_CRITERIA, CRIT_WEIGHTS, WIN_NAME, WIN_ADDRESS, WIN_TOWN, WIN_POSTAL_CODE.

³ Macedonian public procurement law obliges publication of call for tenders above "€500,000 for supplies and services, and €2 million for works in the Official Journal of the EU or in a respective business or technical magazine broadly available to international experts". In line with this obligation, some authorities publish, for free, in the Official Journal, i.e. TED.

period between 2006/01/01 and 2016/12/31. The number of countries covered has increased throughout the years, generally in line with their accession to the European single market.

The data includes selected fields from calls for competition⁵ (most importantly contract notices), contract award notices and voluntary ex-ante transparency notices. The number of fields has been limited (see chapter 4) to avoid having excessively large files with high hardware requirements (e.g. because of free-text descriptions).

- 2.3. Generally, the notices fall under the [EU public procurement directives](#), with the exception of procurement by European institutions, which procure according to the [Financial Regulation](#) of the EU (see CAE_TYPE description in section 4).
- 2.4. Generally, the data consists of notices above the [procurement thresholds](#). However, publishing below threshold notices in TED is considered good practice, and thus a non-negligible number of below threshold notices is present as well.
- 2.5. The data is stored in separate files, with two files per year – one with calls for competition and one with contract award notices – and one file for voluntary ex-ante transparency notices for all years. This separation is done to allow analyses of segments of the data without excessive hardware requirements. For analysis across years or forms, the individual files must be merged.
- 2.6. The data is in comma separated value (CSV) format and is encoded as UTF-8.
- 2.7. Older data has lower quality and sometimes lower coverage, because the data collection structure was less developed. Furthermore, on 17th September 2008, the [common procurement vocabulary was changed](#)⁶. For these reasons, it may often be best to only use data from 2009 onwards. Finally, an important change in data format is taking place throughout 2016 and 2017 and this change may need to be reflected in any analyses. For more information, see the text below and the XSD_VERSION description in section 4.

3. STRUCTURE & INTERPRETATION

- 3.1. The data comes from [public procurement standard forms](#)⁷, which are filled in by contracting bodies and sent as notices for publication to TED. Except a few highlighted exceptions, the variables in the data directly correspond to the fields in the forms. **Before reading further, we strongly recommend reading**

⁴ For technical reasons (different structure of data), notices from the European Investment Bank and the European Bank for Reconstruction and Development are not included in the export.

⁵ A call for competition, in the sense of the directives, is a contract notice, a prior information notice or a periodic information notice used as a call for competition, a concession notice, a public works concession notice (exists only in 2004/18/EC), or a qualification system. Since contract notices represent more than 99 percent of calls for competition and are the most widely known, in this text, we often say contract notices when referring to all of calls for competition.

⁶ There is a many-to-many matching between the old and the new CPV, so they cannot be aligned in a simple manner.

⁷ The standard forms defining the data standard have been established by the [Commission Implementing Regulation \(EU\) 2015/1986](#), [Commission Implementing Regulation \(EU\) No 842/2011](#), [Commission Regulation \(EC\) No 1564/2005](#) and before that by the annexes to the relevant procurement directives.

through a few standard forms (e.g. the most typical ones - the contract notice and the contract award notice). Furthermore, if you are ever unsure how to interpret a variable, your first step should be to find it in the standard forms.

3.2. The data is split into files on contract award notices (**CANs**), contract notices (**CNs**), and voluntary ex-ante transparency notices (**VEATs**). Simply said, a CN informs on a future purchase (“The ministry would like to buy furniture”); a CAN generally informs on the result of the procurement (“The ministry has bought furniture from company X.”). VEATs are likely be interesting only to procurement experts and are explained in the [remedies directives](#)⁸.

3.3. Notices consist of thematic sections. Two of these sections are particularly important, because they have a many-to-one or a many-to-many relationship with other sections, and thus influences the structure of the data. Specifically:

- a single **notice** can contain information about several **lots**,
- a single CAN can contain information about several **contract awards (CA)**,
- a single **contract award** can be linked to several **lots**,
- in some cases⁹, several **contract awards** can be linked to a single **lot**.

3.4. Presenting many-to-one or many-to-many relationship in a flat file format, such as CSV, means that parts of the data will be duplicated within a single file¹⁰.

For example, in a file with CANs, each row begins with information concerning the procedure in general (e.g. the type of procedure), continues with information relevant only for a specific lot, and ends with information relevant only for a specific CA. Since each lot and CA needs its own row, this means that the general part of the information will be repeated on every row.

For example, Table 1 informs about a single CAN informing about three CAs.

Table 1

ID_NOTICE_CAN	description	ID_LOT	lot description	ID_AWARD	value
201501	furniture	ABC	chair	123	€500
201501	furniture	DEF	table	456	€1000
201501	furniture	GHI	cupboard	789	€700

3.5. Which level of data to use depends on the question asked: some fields vary at the notice level (for instance "type of procedure"), some at lot level (for instance the "award criteria" in recent notices), and some at CA level (for instance the "number of bids"). Which field varies at which level depends on where in the form it is located and the version of the form. This can be seen from the form in question, and, for ease of reference, is also listed in the last column of Table 5, below. In general, notice and CA levels are used frequently, lot level information seldom.

⁸ Most points for CANs in the text below are applicable also for VEATs.

⁹ The case of framework agreements, dynamic purchasing systems, innovative partnerships and CANs following qualification systems.

¹⁰ Natively, TED procurement data is stored in XML files and analysed in an SQL database, both of which avoid duplication. Converting this data into a flat file format introduces duplication, but also makes the data more available thanks to the higher number of users who are used to working with CSV files.

For example, if we want to know "How many contract awards in table 1 were related to buying chairs?", then we can simply count the chairs in the "lot description" column of Table 1 and see that there is, indeed, just one. On the other hand, if the question is "How many notices in Table 1 were related to buying furniture?", then before counting anything, we need to remove rows with duplicate observations of ID_NOTICE_CAN. This gives us the correct answer – one.

3.6. Once you know which level of the data you are interested in, you should select it using the "remove duplicates" command, which is available in all statistical as well as spreadsheet programs. The easiest way to remove duplicate rows is to remove them on the basis of duplicate IDs. Table 2 gives an overview of the relevant IDs.

Table 2

Level	unique identifier	Notes
notice (also called "procedure")	ID_NOTICE_CAN, ID_NOTICE_CN, etc.	Identifier for the entire notice.
lot	ID_LOT	Identifier for section II.2 of a notice. This is available only for 2.0.9 forms ¹¹ . Theoretically, this lot identifier should be unique within a notice. Regrettably, in practice this is not yet always the case because of technical errors.
contract award	ID_AWARD	Identifier for section V of a CAN or a VEAT.

3.7. As mentioned in section 3.3, the relationship between lots and CA can be more complicated than the one described in Table 1. Furthermore, because of missing validation rules, it can be impossible to match contracts and lots – in which case the data will look like Table 3 instead of Table 1.

Table 3

ID_NOTICE_CAN	description	ID_LOT	lot description	ID_AWARD	value
201501	furniture	ABC	chair		
201501	furniture	DEF	table		
201501	furniture	GHI	cupboard		
201501	furniture			123	€500
201501	furniture			456	€1000
201501	furniture			789	€700

Nevertheless, note that even in this case, all that is needed is to remove duplicates on a particular identifier, and the data will have the right form to answer relevant questions.

¹¹ In previous forms, lot information was included in Annex B, but it is not included in the data because of very low quality.

3.8. Finally, note that the same type of duplications as between CANs, lots, and CAs will also occur when merging data from CAN and CN level (which is done by using FUTURE_CAN_ID). For example, if the CAN in Table 1 was preceded by two CNs, the merged database would have the following structure:

Table 4

ID_NOTICE_CAN	description	ID_LOT	lot description	ID_AWARD	description	ID_NOTICE_CN
201501	furniture	ABC	chair	123	chair	201442
201501	furniture	DEF	table	456	table	201442
201501	furniture	GHI	cupboard	789	cupboard	201442
201501	furniture	ABC	chair	123	chair	201466
201501	furniture	DEF	Table	456	table	201466
201501	furniture	GHI	Cupboard	789	cupboard	201466

3.9. The data includes also notices which were published and then cancelled. This is indicated by the value of CANCELLED being “1”. For most types of analyses, you will probably want to drop these notices. No information about cancellation is available for notices published before 2011.

Understanding the structure of the data is crucial. If you do not feel 100% at ease with the explanations above, please review them once more while looking at the data.

4. CODEBOOK

This chapter gives basic descriptions of the available variables. The first column provides variable names. The second and third column their description for CFCs and CANs/VEATs, respectively. If the description is the same for both types of notices, it is included in one cell. The fourth column explains which level (see above) the variable is at, whether this has changed in different versions of the forms, and whether the variable is available in all versions of the forms.

In case of doubt about what a variable means, please refer to the [public procurement standard forms](#). **Quotation marks in the description columns indicate direct quotations from the standard forms.** Please note that the exact names from the standard forms might have been slightly different in different versions of the forms.

When a variable does not directly correspond to a field in the standard forms, but was added to the notice on the basis of our calculations, we mention it was "[ADDED]" in the Description column.

Table 5

Variable	CFC description	CAN/VEAT description	Level (since which version)
Notice metadata			
ID_NOTICE_CN	<p>Unique identifier of the call for competition (usually contract notice).</p> <p>This is the identifier for all variables at the notice level.</p>		notice
ID_NOTICE_CAN / ID_NOTICE_VEAT		<p>Unique identifier of the contract award notice / voluntary ex-ante transparency notice.</p> <p>This is the identifier for all variables at the notice level.</p>	notice

TED_NOTICE_URL	Webpage of the notice on the TED website. Having a look can give a more qualitative insight into what is being procured. Note that TED hosts notices only for five years after publication, so for notices older than that the link will not work.	notice
YEAR	Year of publication of the notice	notice
ID_TYPE	Standard form number, see the relevant TED webpage .	notice
DIRECTIVE	The VEAT standard form can be used under several directives. This variable specifies the directive. (For other types of notices, the directive type is based on ID_TYPE.)	notice
DT_DISPATCH	The date when the buyer dispatched (sent) the notice for publication to TED.	notice
XSD_VERSION	Version of the XML schema definition used by the Publications Office of the EU to publish the data. Higher versions mean better average quality of data. The lowest version is "2.0.5", the highest "2.0.9"; notices before 2006 are of the lowest quality and do not have version information. [ADDED]	notice
CANCELLED	1 = this notice was later cancelled [ADDED]	notice
CORRECTIONS	Number of later notices which corrected or added information to this notice (see standard form 14). [ADDED]	notice

FUTURE_CAN_ID¹²	The publication ID of the CAN which followed this notice. This ID is used to link the CFC and CAN datasets (by putting the FUTURE_CAN_ID equal to the ID_NOTICE_CAN). [ADDED]		notice
FUTURE_CAN_ID_ESTIMATED	Whether the "future" publication ID submitted in the notice was estimated (corrected) for this dataset, for instance because of a straightforward typo. This variable can explain differences compared to the TED website. 1 = estimated. [ADDED]		notice
Contracting authority or entity identification			

¹² This variable is actually based on a link made in a CAN (section IV.3.2), where CFCs are indicated as "previous publications". For technical reasons, we provide this information in the opposite manner, i.e. as "future publication" of a CN. Please note that the values in this variable reflect any link made between a CN and CAN, no matter in which section of the "previous publication" it was included. On the other hand, links to other types of notices (often accidental) are excluded.

B_MUTIPLE_CAE	<p>There is more than one contracting authority or entity.</p> <p>If this is the case, each row below in this section (with the exception of ISO_COUNTRY_CODE) will contain information per each authority, separated by "---". ISO_COUNTRY_CODE will contain only the information for the first listed authority. [ADDED]</p>	notice (only in XSD_VERSION = 2.0.9)
CAE_NAME	"Official name"	notice
CAE_NATIONALID	"National registration number" e.g. VAT number for utilities	notice
CAE_ADDRESS	"Postal address"	notice
CAE_TOWN	"Town"	notice
CAE_POSTAL_CODE	"Postal code"	notice
ISO_COUNTRY_CODE	"Country" for the first listed authority	notice
B_MULTIPLE_COUNTRY	There are contracting authorities or entities from at least two different countries. [ADDED].	notice (only in XSD_VERSION = 2.0.9)
ISO_COUNTRY_CODE_ALL	If the variable above is yes, then this variable contains the list of all countries.	notice (only in XSD_VERSION = 2.0.9)
Other notice level and lot level variables		

CAE_TYPE

Type of contracting authority.
1 "Ministry or any other national or federal authority, including their regional or local subdivisions"
3 "Regional or local authority"
4 "Water, energy, transport and telecommunications sectors"
5 "European Union institution/agency"
5A "other international organisation"
6 "Body governed by public law"
8 "Other"
N "National or federal Agency / Office"
R "Regional or local Agency / Office"
Z "Not specified"

The distinction between 5 and 5A has been [ADDED] on the basis of data not included in the standard forms.

Please note that procurement by "European Union institution/agency" will generally not be covered by public procurement legislation, but by the [Financial Regulation of the EU](#). Thus, it may be appropriate to exclude them from analyses dealing with the procurement directives. Similarly, it might be appropriate to exclude these observations for analyses of national level procurement, since the responsibility for this procurement lies at the EU level.

notice

EU_INST_CODE	<p>EU institution (or type of EU institution).</p> <p>AG "agencies" BC "European Central Bank" BI "European Investment Bank" BR "European Bank for Reconstruction and Development" CA "European Court of Auditors" CJ "Court of Justice of the European Union" CL "Council of the European Union" CR "European Committee of the Regions" EA "European External Action Service" EC "European Commission" ES "European Economic and Social Committee" FI "European Investment Fund" OB "European Patent Office" OP "Publications office of the European Union" PA "European Parliament"</p> <p>If CAE_TYPE is not 5, then this variable is empty.</p>	notice (only in XSD_VERSION = 2.0.9)
MAIN_ACTIVITY	<p>(The classification corresponds to COFOG divisions.)</p> <p>In XSD_VERSION = 2.0.9 this variable newly contains exactly one value.</p>	notice
B_ON_BEHALF	This indicates either a central purchasing body or several buyers buying together (i.e. occasional joint procurement).	notice
B_INVOLVES_JOINT_PROCUREMENT	"The contract involves joint procurement"	notice (only in XSD_VERSION = 2.0.9)
B_AWARDED_BY_CENTRAL_BODY	"The contract is awarded by a central purchasing body"	notice (only in XSD_VERSION = 2.0.9)
TYPE_OF_CONTRACT	<p>Type of contract. The values are the following:</p> <p>W "Works" U "Supplies" S "Services"</p>	notice

TAL_LOCATION_NUTS	The Nomenclature of Territorial Units for Statistics (NUTS) code placement of the "Main site or location of work, place of delivery or of performance"	notice
B_FRA_AGREEMENT	Y if "The notice involves the establishment of a framework agreement" is selected or "Framework agreement with a single operator" is selected or "Framework agreement with several operators" is selected.	"The notice involves the establishment of a framework agreement"
		notice

FRA_ESTIMATED

Whether there are indications that this notice is actually about a framework agreement, even though it has not been marked as such by the buyer (i.e. the buyer possibly forgot to mark the field). Indications are the following:

K "The keyword `framework', in the appropriate language, was found in the title or description of the notice."

C "Consistency across notices: at least half of the contract award notices which followed this notice were marked as framework agreement

Whether there are indications that this notice is actually about a framework agreement, even though it has not been marked as such by the buyer (i.e. the buyer possibly forgot to mark the field). Indications are the following:

K "The keyword `framework', in the appropriate language, was found in the title or description of the notice."

A "Multiple awards were given per one lot, which is legally admissible only in case of framework agreements, dynamic purchasing systems, innovation partnerships, and qualification systems."

C "Consistency across notices: the contract notice which preceded this notice was marked as a framework agreement."

The use of these indications depends on how important it is to not misclassify frameworks in your analysis. One possible approach is to assume that the notice has been misclassified as a framework when at least two of the indications above are present. [ADDED]

notice

B_FRA_CONTRACT		This notice is probably about specific contracts within a framework agreement. This has been estimated on the basis of the previous notice being a contract award notice and announcing the establishment of a framework agreement. [ADDED]	notice
B_FRA_SINGLE_OPERATOR	Y "Framework agreement with a single operator" N "Framework agreement with several operators"		notice
FRA_NUMBER_OPERATORS	"Number of participants to the framework agreement envisaged"		notice
FRA_NUMBER_MAX_OPERATORS	"Maximum number of participants to the framework agreement envisaged"		notice
B_DYN_PURCH_SYST	"The notice involves contract(s) based on a dynamic purchasing system"		notice
CPV	The main Common Procurement Vocabulary code of the main object of the contract		notice
ID_LOT	lot identifier for the section II.1 of the notice This is the identifier for all variables at the lot level.		lot (only in XSD_VERSION = 2.0.9)
ADDITIONAL_CPV	For XSD<2.0.9, this is the main Common Procurement Vocabulary of the additional object(s) of the contract. For XSD=2.0.9, these are the CPV codes of the lots.		notice in XSD_VERSION < 2.0.9 lot in XSD_VERSION = 2.0.9
B_GPA	"The contract is covered by the Government Procurement Agreement"		notice

LOTS_NUMBER	<p>The number of lots for a given call for competition.</p> <p>From a legal point of view, not dividing a notice means may be interpreted as not using lots, and thus in these cases the variable will have value of 0. Value of 1 is a user imputed mistake. This value is based on the number of Annexes B of a CN. [ADDED]</p>	<p>The number of lots for a given CAN. This value is based on the number of unique "Lot No" values included in the heading of section V of the CAN. This information is available only since 2009. [ADDED]</p>	notice
LOTS_SUBMISSION	<p>"Tenders may be submitted for:"</p> <ul style="list-style-type: none"> - A "all lots" - O "one lot" - M "multiple lots" 	/	Notice
B_VARIANTS	<p>"Variants will be accepted"</p>	/	notice in XSD_VERSION < 2.0.9 lot in XSD_VERSION = 2.0.9
VALUE_EURO	<p>CFC value, in EUR, without VAT. If the value was not present, the lowest part of the range is used instead.</p>	<p>CAN value, in EUR, without VAT. If the value was not present, the lowest bid is used instead.</p>	notice

VALUE_EURO_FIN_1	CFC value, in EUR, without VAT. If the value variable is missing, the framework value is used instead (see explanation in Annex I.). [ADDED]	CAN value, in EUR, without VAT. If a value variable is missing, this variable looks for it in all other fields from which it could be taken (see detailed explanation in Annex I). [ADDED]	notice
VALUE_EURO_FIN_2	Generally the same value as VALUE_EURO_FIN_1, but – if available – overwritten by human-made estimates of values for large value contracts which seemed to be incorrect. ¹³ These estimates are indicative and the Commission holds no responsibility for their correctness and use. [ADDED]		notice
B_OPTIONS	"Options"		notice in XSD_VERSION < 2.0.9 lot in XSD_VERSION = 2.0.9
B_EU_FUNDS	"The contract is related to a project and / or programme financed by European Union funds"	"The contract is related to a project and / or programme financed by European Union funds"	notice in XSD_VERSION < 2.0.9 lot in XSD_VERSION = 2.0.9
B_RENEWALS	"This contract is subject to renewal"		notice
DURATION	"Duration of the contract, framework agreement or dynamic purchasing system" in months (see explanation in Annex II).		notice in XSD_VERSION < 2.0.9 lot in XSD_VERSION = 2.0.9

¹³ As explained in presentations which are part of this documentation, typos in large value contracts can severely bias aggregate statistics even at national level. For this reason, the value of some large value contracts has been manually corrected throughout the years, e.g. on the basis on the information available in the notice description, contacting the buyer, or searching online.

CONTRACT_START	"Duration of the contract or time limit for completion"; "Starting"		notice in XSD_VERSION < 2.0.9 lot in XSD_VERSION = 2.0.9
CONTRACT_COMPLETION	"Duration of the contract or time limit for completion"; "Completion"		notice in XSD_VERSION < 2.0.9 lot in XSD_VERSION = 2.0.9
TOP_TYPE	Type of procedure. The values are the following: AWP "award without prior publication of a contract notice" COD "competitive dialogue" NEC/NEG "negotiated with a call for competition" OPE "open" RES "restricted"	Type of procedure. The values are the following: AWP "award without prior publication of a contract notice" COD "competitive dialogue" NOC/NOP "negotiated without a call for competition" NIC/NIP "negotiated with a call for competition" OPE "open" RES "restricted"	notice
B_ACCELERATED	The option to accelerate the procedure has been used. This is possible for negotiated, restricted, and (under the 2014 directives) open procedures.		notice
OUT_OF_DIRECTIVES		"The procurement falls outside the scope of application of the directive", but a CAN was published anyway. (See CAN Annex D1). This field exists only when CANs are published without a CN, i.e. when TOP_TYPE is NOC/NOP or AWP.	notice
ENV_OPERATORS	"Envisaged number of operators"		notice
ENV_MIN_OPERATORS	"Envisaged minimum number"		notice
ENV_MAX_OPERATORS	"and (if applicable) maximum number"		notice

CRIT_CODE	Award criteria. The values are the following: L "Lowest price" M "Most economically advantageous tender"	notice in XSD_VERSION < 2.0.9 lot in XSD_VERSION = 2.0.9
CRIT_PRICE_WEIGHT	Weight given to price. [Added]	lot (only in XSD_VERSION = 2.0.9)
CRIT_CRITERIA	Information on award criteria. For XSD<2.0.9, the variable is usually unstructured text. For XSD=2.0.9, the variable contains all criteria names separated by "---", except for price, because this is already in CRIT_PRICE_WEIGHT.	notice in XSD_VERSION < 2.0.9 lot in XSD_VERSION = 2.0.9
CRIT_WEIGHTS	Information on award criteria weighing. For XSD<2.0.9, the variable is usually an unstructured text field. For XSD=2.0.9, the variable contains all criteria names separated by "---", except for price, because this is already in CRIT_PRICE_WEIGHT.	notice in XSD_VERSION < 2.0.9 lot in XSD_VERSION = 2.0.9
B_ELECTRONIC_AUCTION	"An electronic auction has been used"	notice
NUMBER_AWARDS		The number of CAs for a given CAN. [ADDED]
DT_APPLICATIONS	"Time limit for receipt of tenders or requests to participate"	
B_LANGUAGE_ANY_EC	"Language(s) in which tenders or requests to participate may be drawn up - Any EU official language"	
ADMIN_LANGUAGES_TENDER	"Language(s) in which tenders or requests to participate may be drawn up - Official EU language(s):"	
ADMIN_OTHER_LANGUAGES_TENDER	"Language(s) in which tenders or requests to participate may be drawn up - Other:"	

B_RECURRENT_PROCUREMENT	"This is a recurrent procurement"		notice
Award metadata			
ID_AWARD		unique contract award identifier This is the identifier for all variables at the contract award level.	contract award
ID_LOT_AWARDED		"Lot No", an identifier of a lot within this contract award In theory, this identifier should be the same as one of the section II.1 lot identifiers. In reality, in some cases they may differ because of missing validation rules.	contract award
INFO_ON_NON_AWARD		If the variable is empty, then a contract was awarded. PROCUREMENT_UNSUCCESSFUL means that "A contract is not awarded", because "No tenders or requests to participate were received or all were rejected" PROCUREMENT_DISCONTINUED means that "A contract is not awarded" because of "Other reasons (discontinuation of procedure)"	contract award
Winning bidder identification			
B_AWARDED_TO_A_GROUP		"The contract has been awarded to a group of economic operators" Currently, only information about the economic operator listed first is available in the data.	contract award (only in XSD_VERSION = 2.0.9)
WIN_NAME		"Official name"	contract award

WIN_NATIONALID		"National registration number" e.g. a VAT number of a business registry number	contract award (only in XSD_VERSION = 2.0.9)
WIN_ADDRESS		"Postal address"	contract award
WIN_TOWN		"Town"	contract award
WIN_POSTAL_CODE		"Postal code"	contract award
WIN_COUNTRY_CODE		"Country"	contract award
B_CONTRACTOR_SME		"The contractor is an SME (SME as defined in Commission Recommendation 2003/361/EC)"	contract award (only in XSD_VERSION = 2.0.9)
Other CA level variables			
CONTRACT_NUMBER		"Contract No"	contract award
TITLE		"Title"	contract award
NUMBER_OFFERS		"Number of tenders received"	contract award
NUMBER_TENDERS_SME		"Number of tenders received from SMEs (SME as defined in Commission Recommendation 2003/361/EC)"	contract award (only in XSD_VERSION = 2.0.9)
NUMBER_TENDERS_OTHER_EU		"Number of tenders received from tenderers from other EU Member States"	contract award (only in XSD_VERSION = 2.0.9)
NUMBER_TENDERS_NON_EU		"Number of tenders received from tenderers from non-EU countries"	contract award (only in XSD_VERSION = 2.0.9)
NUMBER_OFFERS_ELECTR		"Number of offers received by electronic means"	contract award
AWARD_EST_VALUE_EURO		Estimated CA value, in EUR, without VAT.	contract award
AWARD_VALUE_EURO		Total final CA value, in EUR, without VAT. If the value was not present, the lowest bid is included.	contract award
AWARD_VALUE_EURO_FIN_1		CA value, in EUR, without VAT. If the value variable is missing, this variable looks for it in all other fields from which it could be taken (see detailed explanation in Annex I). [ADDED]	contract award
B_SUBCONTRACTED		"The contract is likely to be subcontracted."	contract award
DT_AWARD		"Date of contract award"	contract award

5. ANNEX I – CALCULATION OF VALUE_EURO_FIN_1 (and AWARD_VALUE_EURO_FIN_1)

(Note: all values below are taken in EUR and without VAT.¹⁴)

1. We take the sum of awards, which sums up:
 - 1.1. the total final value of the award V.2.4 (V.4¹⁵ in previous standard form version 2.0.8),
 - 1.2. if a value of a given award is not available, then the value of the lowest offer V.2.4 (V.4 in previous standard form version 2.0.8),
 - 1.3. if neither of these is available, then the highest offer V.2.4 (V.4 in previous standard form version 2.0.8),
 - 1.4. if the above is missing,
 - A) but the section “Official name” in section V.2.3. is not empty (as this may indicate that no winner was chosen for this award), the “initial estimated value of the award” V.2.4 (V.4 in previous standard form version 2.0.8), is used instead.
 - B) and “Official name” in section V.2.3. is also empty, this value is treated as 0 for the calculation of the sum.

(The value resulting from steps 1.1 – 1.3 is stored in AWARD.VALUE_EURO_FIN_1.)

2. If the above is not available, or the notice is a framework agreement or a dynamic purchasing system (or if this information is missing but there are several contract awards for one lot) and the sum of awards above is higher than value in the Object of the Contract II.1.7 (II.2.1 in previous standard form version 2.0.8)¹⁶, or the sum of awards – including multiplication by number of months or years valid only for previous standard form version 2.0.8 – is the same as the Object of the Contract value, or there are two or more award values which are exactly the same as the Object of the Contract value¹⁷, we take the Object of the Contract value of the notice:
 - 2.1. The value of the contract II.1.7 (II.2.1 in previous standard form version 2.0.8),
 - 2.2. if it is not available, then the value of the lowest offer II.1.7 (II.2.1 in previous standard form version 2.0.8),
 - 2.3. if neither of these is available, then highest offer II.1.7 (II.2.1 in previous standard form version 2.0.8).

¹⁴ Valid only for previous standard forms, xsd version 2.0.8, - If information is missing on whether a value includes VAT or not, the information is taken from other fields in the same form (i.e. II.2.1 or V.4). If all fields are missing, we assume that VAT is excluded.

¹⁵ If annual or monthly values are given (see bottom of field V.4), they are ignored (unless they are the same as the object of the contract value, see point 2). This approach has been chosen because while the forms clearly say that the number of months/years should be given only when a monthly/annual value was submitted, in reality it seems that users tend to submit the duration of their contract into this field regardless of whether the value was monthly/yearly.

¹⁶ This condition is included because framework agreements and dynamic purchasing systems may have several contract awards for one lot, in which case summing up the contract awards overestimates the value. The second part of the condition is added to cover cases when a CAN is used to report on new contracts awarded *within* an already existing framework agreement and to keep the estimate conservative.

¹⁷ This condition may indicate that a notice value was mistaken for an award value and included multiple times.

3. If the above is not available, the CAN has at least one non-empty “Official name” in section V.2.3., and all of the CNs in the “previous publication” (IV.2.1) of this CAN are followed only by this one CAN, then the best value for the CN¹⁸ is taken, being:

3.1. The first existing value (estimated, minimal, maximal) in “Estimated value of the contract notice” II.1.5 (II.2.1 in previous standard form version 2.0.8),

3.2. Only for previous forms version 2.0.8 - If none of these values are available, the first existing value (estimated, minimal, maximal) “Estimated total value of purchases for the entire duration of the framework agreement” (section II.1.4) in the CN is taken.

When there were multiple CNs, then the value of the lowest one is taken¹⁹.

Reminder: As discussed in point 2.1, the quality of (value) data submitted by contracting authorities is often low. In most cases, the rules listed above are not sufficient for relying on the values of the notices without further quality checks.

¹⁸ Or a prior or periodic information notice used as a call for competition.

¹⁹ Multiple CNs are caused most often by a buyer accidentally uploading two versions of the same notice. Thus, the sum of all CNs is not an appropriate estimate. The lowest value is taken as a conservative estimate.

6. ANNEX II – CALCULATION OF DURATION

1. Duration in months (CN II.3),
2. If not available, then the duration in days divided by the average length of the month (30.4) (CN II.3).
3. If not available, then the period between the day of starting and day of completion of the contract divided by the average length of the month (30.4) (CN II.3).
4. If not available, then the duration of the framework agreement (CN II.1.4) in months.
5. If not available, then the duration of the framework agreement (CN II.1.4) in years multiplied by the number of months (12).
6. If not available, then the latest date of completion of a lot minus the earliest day of starting of a lot, divided by the average length of the month (30.4) (CN Annex B, 4)).

The resulting values are rounded to two decimal places.

7. ANNEX III – VERSION HISTORY

Updates in version 3.1 (2018-02-23)

- Addition of B_FRA_CONTRACT.
- Correction of a significant amount of missing data for the CN 2016 file.
- Correction of award criteria information under XSD_VERSION 209.
- Correction of a very limited number of errors, concerning both missing / duplicated notices and incorrect variable values.
- Updated description of Annex V - Calculation of Value_euro_fin_1 in this documentation.

Updates in version 3.0 (2017-08-02)

- Addition of 2016 data. Roughly 1/3 of notices in 2016 are in the new XSD_TYPE 2.0.9 format. This brings several changes to the data, explained throughout this documentation. In particular:
 - Many more variables can now change per lot (see chapter 4 for details). Thus, information about lots has been added to the notices. ID_LOT is the identifier used in section II of notices.
 - To avoid confusion, the lot identifier used in section V of CANs/VEATs has been renamed from LOT_NUMBER to ID_LOT_AWARDED.
 - Addition of variables newly available in the 2.0.9 format (B_INVOLVES_JOINT_PROCUREMENT, B_AWARDED_BY_CENTRAL_BODY, B_MULTIPLE_CAE, B_MULTIPLE_COUNTRY, ISO_COUNTRY_CODE_ALL, CRIT_PRICE_WEIGHT, B_ACCELERATED, B_AWARDED_TO_A_GROUP, WIN_NATIONALID, B_CONTRACTOR_SME, NUMBER_TENDERS_SME, NUMBER_TENDERS_OTHER_EU, NUMBER_TENDERS_NON_EU). See chapter 4 for details.
 - Accelerated procedures have been removed from TOP_TYPE. Instead, a binary variable B_ACCELERATED has been added.
- Addition of other variables (EU_INST_CODE, OUT_OF_DIRECTIVES TED_NOTICE_URL, DIRECTIVE). See chapter 4 for details.
- Voluntary ex-ante transparency notices have been added to the export.
- Extending the coverage of the "contract notice" files, so that they include all types of calls for competition. A call for competition, in the sense of the directives, is a contract notice, a prior information notice or a periodic information notice used as a call for competition, a concession notice, a public works concession notice (exists only in 2004/18/EC), or a qualification system.
- Tens of notices, which were missing because of a technical error, have been added for 2015 together with a few 2.0.9 notices for 2015.
- Correction of a technical error because of which EU institutions were not included in the calls for competition export.
- In CFC files, the PROCEDURE variable has been renamed to TOP_TYPE to be named the same as in the CAN files.

- ADDITIONAL_CPV1_4 has been renamed to ADDITIONAL_CPV, as it contains all CPV values, not just the first four.
- Missing values in B_FRA_AGREEMENT in CANs have been recoded to "N", so as to be consistent with the notation used in CNs.
- In case of multiple titles in CAN section V (the variable TITLE), these titles have been put in the same order as is used in LOT_NUMBER.
- In all variables which concatenate multiple strings, these strings are now always separated by "---". This means the delimiter has changed for CRIT_CRITERIA and ADDITIONAL_CPV.
- Codes used in XSD_TYPE have been shortened.

Updates in version 2.2 (2016-09-15)

- Correction of a technical error because of which FRA_AGREEMENT were empty for contract notices.
- Addition of a new variable FRA_ESTIMATED.
- Addition of new variables: B_FRA_SINGLE_OPERATOR, FRA_NUMBER_OPERATORS, FRA_NUMBER_MAX_OPERATORS for contract notices.
- B_FRA_AGREEMENT for contract notices is now marked as "Y" also when a value is present in B_FRA_SINGLE_OPERATOR.
- Minor change in the calculation of "VALUE_EURO_FIN_1": when taking the lower of the values (e.g. for "lowest offer"), it now takes the objectively smaller one, not the one which was entered by the contracting authority in the "smaller value" field.
- Minor change in the calculation of "VALUE_EURO_FIN_1": when looking for values in contract notices, periodic information notices (SF4) used as a call for competition are also taken into consideration.
- Addition of several manual corrections for "VALUE_EURO_FIN_2" for France in 2013.
- Ensuring that the variable names CORRECTIONS, ADMIN_LANGUAGES_TENDER, and ADMIN_OTHER_LANGUAGES_TENDER are consistently used in all exports.

Updates in version 2.1 (2016-07-01)

- Correction of a technical error on the basis of which notices towards the end of a given year had a missing value for the VALUE_EURO_FIN_1 and VALUE_EURO_FIN_2 variables.
- Correction of a technical error where notices which were cancelled were included twice in the export – once marked as cancelled, once as not cancelled. Please note that cancellation is reliably available only for notices from 2011 onwards.
- Addition of a new condition in the methodology for VALUE_EURO_FIN_1 and VALUE_EURO_FIN_2 for frameworks and dynamic purchasing systems and of explicit assumptions on whether values include VAT or not (see Annex I).
- Addition of manual corrections for VALUE_EURO_FIN_2 for 2015.
- Addition of a new variable CORRECTIONS.

- Addition of a new source of data for DURATION in contract notices, see Annex 2.
- Addition of number of lots for contract award notices. Available only since the end of 2009.
- Addition of a new variable XSD_TYPE.
- Addition of notices published by EU institutions. While these notices are generally covered by the financial regulation of the EU, not by the procurement directives, they use the same standard forms. Furthermore, the category of “EU institution/agency of international organisation” has been split into two: “EU institution/agency” and “other international organisation”.
- Various minor updates in the documentation, including the addition of a version history and a list of file sizes to allow checking whether the files were correctly downloaded.
- VALUE_EURO_FIN_1_1 was renamed to AWARD_VALUE_EURO_FIN_1 in the export, as used in the documentation.

Updates in Version 2.0 (2016-04-12)

- The export has been divided into files per year and type of notice.
- Contract notices have been added to the export.

8. ANNEX IV – DOWNLOAD SIZES

The table below allows checking whether the files were downloaded correctly.

File name	Size (KB)
TED_CAN_2006.csv	134,804
TED_CAN_2007.csv	176,160
TED_CAN_2008.csv	204,497
TED_CAN_2009.csv	227,532
TED_CAN_2010.csv	265,490
TED_CAN_2011.csv	289,684
TED_CAN_2012.csv	306,788
TED_CAN_2013.csv	309,948
TED_CAN_2014.csv	326,109
TED_CAN_2015.csv	341,990
TED_CAN_2016.csv	335,755
TED_CAN_2017.csv	439,822
TED_CAN_2009_2017.csv	2,834,104
TED_CN_2006.csv	62,937
TED_CN_2007.csv	68,756
TED_CN_2008.csv	73,667
TED_CN_2009.csv	76,032
TED_CN_2010.csv	80,055
TED_CN_2011.csv	83,303
TED_CN_2012.csv	83,257
TED_CN_2013.csv	86,267
TED_CN_2014.csv	86,535
TED_CN_2015.csv	90,347
TED_CN_2016.csv	366,632
TED_CN_2017.csv	522,048
TED_CN_2009_2017.csv	1,504,464
TED_VEAT	71,973