

Codebook v 1.01

The IUROPA CJEU Database Platform

By Joshua C. Fjelstul, Ph.D.

Contents

The IUROPA CJEU Database Platform	2
Cases	8
Proceedings	13
Decisions	19
Judges	25
Judgments	29
Opinions	39
Appointments	44
Parties	47
Procedures	51
Assignments	57
Observers	63
Citations	68
Actors (summary)	74
Formations (summary)	76
Procedures (summary)	78
Document types (summary)	80
Proceeding suffixes (summary)	82

The IUROPA CJEU Database Platform

A component of the IUROPA CJEU Database

Introduction

The IUROPA CJEU Database Platform, published by the IUROPA Project, is the largest and most comprehensive research-oriented database on the Court of Justice of the European Union (CJEU). It includes data on the universe of CJEU cases, proceedings, decisions, and judges collected from the official Registry of the Court, InfoCuria (the Court's official database), and EUR-Lex (the EU's official database of legal documents). All of the data has been cleaned and cross-validated and is research-ready. This codebook describes the tables and variables in the IUROPA CJEU Database Platform and provides important data notes for researchers.

We start by creating a complete list of all cases, proceedings, decisions (e.g., judgments, orders, Advocate General opinions, etc.), and judges (including Advocates General). These are the 4 basic units of analysis in the database. We create one table for each unit of analysis: `cases`, `proceedings`, `decisions`, and `judges`. The `decisions` table has 2 sub-tables, `judgments` and `opinions`, that include additional variables that are specific to those types of decisions.

Then, we create 6 tables with more complex units of observation: `appointments` (one observation per appointment per judge), `parties` (one observation per party per proceeding), `procedures` (one observation per procedure per decision), `assignments` (one observation per judge per decision), `observers` (one observation per institutional actor that submits an observation with respect to a judgment), and `citations` (one observation per unique citation per decision).

We also provide **5** supplemental tables: `summary_actors`, which includes information about the various institutional actors that appear in the dataset (e.g., EU member states, EU institutions, EU agencies, etc.); `summary_formation`, which includes information about the formations of the Court (e.g., Full Court, Grand Chamber, First Chamber, etc.); `summary_procedures`, which includes information about the various legal procedures at the Court (e.g., references for a preliminary ruling, actions for annulment, applications for interim measures, etc.); `summary_proceeding_suffixes`, which includes information about proceeding suffixes (i.e., suffixes for case IDs that the Court uses to differentiate between proceedings associated with the same case); and `summary_document_types`, which includes information about the types of documents cited in CJEU decisions.

The IUROPA CJEU Database Platform addresses **3** fundamental problems with data on CJEU decisions from official sources to make quantitative analysis of CJEU decisions easier: (1) there are inconsistencies across official data sources with respect to basic information like decision IDs, case IDs, proceeding IDs, and decision dates, (2) ECLI numbers do not always uniquely identify decisions, and (3) it is difficult to accurately identify exactly which cases and proceedings each decision applies to because of how the Court joins cases together.

There are **2** general problems with ECLI numbers: (1) sometimes one ECLI number is assigned to more than one decision, and (2) sometimes more than one ECLI number is assigned to one decision. By cross-referencing data from the Registry, InfoCuria, and EUR-Lex, we address both of these problems and are able to construct a complete table of unique decisions. We provide a unique IUROPA decision ID for each decision that indicates the case of record, the type of the decision, and the date of the decision. These are more informative than ECLI numbers, which do not indicate the type or date of a decision and cannot be constructed using information about a decision.

The tables

The IUROPA CJEU Database Platform includes **4** core tables: `cases`, `proceedings`, `decisions`, and `judges`. These tables cover the basic units of analysis in the database. The `decisions` table has two sub-tables: `judgments` and `opinions`,

which include variables that are relevant to judgments and AG opinions, the two types of decisions most researchers will be interested in. In future updates, we plan to include additional data on orders.

The database also includes 6 tables with more complex units of observation: `appointments` (one observation per appointment per judge), `parties` (one observation per party per proceeding), `procedures` (one observation per procedure per decision), `assignments` (one observation per judge per decision), `observers` (one observation per institutional actor that submits an observation with respect to a judgment), and `citations` (one observation per unique citation per decision).

The `appointments` table uses judges as the base unit of observation. Many individuals who serve as judges at the CJEU serve in multiple positions. Many serve on one Court and then move to another. Others serve as an AG and then as a judge. In addition, some judges become President or (at the Court of Justice) Vice-President. This table has one observation per appointment per judge, with information about each position an individual has held at the Court. This table also includes the registrars.

The `parties` table uses proceedings as the base unit of observation. When multiple cases are joined into a single proceeding, the parties still vary across the joined proceedings. In addition, some cases involve multiple legal procedures that necessitate multiple proceedings (e.g., applications as to costs, applications for interim measures, etc.). The parties can vary by proceeding, even within a case.

The `procedures`, `assignments`, `observers`, and `citations` tables each use decisions as the base unit of observation. Citations are clearly decision-specific, and observers are only relevant to judgments concerning references for a preliminary ruling, but it is less obvious whether the `procedures` and `assignments` tables should use decisions or proceedings as the base unit of observation. We use decisions as the base unit of observation for the `procedures` and `assignments` because the legal procedure and the judges who wrote the decision are both reported in the text of the decision. It is possible for the composition of the panel to change between judgments if a judge becomes unavailable and has to be replaced.

The `procedures` and `assignments` tables currently only cover judgments. The `observers` table covers judgments concerning references for preliminary rulings at the

Court of Justice. The `citations` table covers all decisions that EUR-Lex has citation metadata for (this EUR-Lex metadata the source of the data). In future updates, we plan to collect additional data on citations to make this table as complete as possible.

Cases versus proceedings

It is important to distinguish between cases and proceedings. When a case is lodged, the Court assigns a case number (e.g., C-370/12). In the database, the variable `cjeu_case_id` records this value. A `cjeu_case_id` indicates the Court at which the case was lodged, the year of the case, and the number of the case (within the year). The `cases` table includes every `cjeu_case_id` that appears in the Registry plus all `cjeu_case_id` that should exist, which is any `cjeu_case_id` with a `case_number` between `1` and the highest value of `case_number` that occurs in the Registry each year.

A proceeding is an action before the Court involving a specific set of cases and legal procedures. The `proceedings` table includes every unique proceeding that appears in the Registry (what it has its own record or is mentioned in the record for another proceeding) or that is associated with a CJEU decision. A case always has a main proceeding (e.g., an action for annulment, a reference for a preliminary ruling, etc.), and that proceeding can include multiple legal procedures (e.g., an action for annulment and an action for damages).

Generally, the main proceeding involves one of the major procedures outlined in Section 5 of the Treaty on the Functioning of the European Union (TFEU), such as a reference for a preliminary ruling, a direct action (i.e., an action for annulment, an action for failure to act, an action for failure to fulfill obligations, or an action for damages), or a staff case. But it is also possible for a main proceeding to involve a minor procedure, such as an application for legal aid. Rarely, there is a record of what would usually be a secondary proceeding, such as a proceeding as to costs, but no record of a main proceeding.

A case can also have additional, secondary proceedings. Some procedures require a separate proceeding. Some examples of secondary proceedings are applications as to costs, applications for interim measures, applications for legal aid, and applications for revision of a decision. In this sense, a proceeding is narrower than a case. The Court can also join proceedings together, in which case a proceeding could involve multiple cases. In

this sense, a proceeding is broader than a case. A secondary proceeding can also involve joined proceedings, in which case the proceeding is narrower than a case in the first sense, but also broader in the second.

Importantly, for the purposes of the `parties` table, the parties in secondary proceedings can be different than the parties in the main proceedings, as an application can be submitted by a subset of applicants or by one or more defendants. This is why the base unit of observation for the parties table is a proceeding, rather than a case.

Proceedings are uniquely identified by a proceeding ID, which is recorded in the database by the variable `cjeu_proceeding_id`. A proceeding ID is a case ID that is sometimes followed by a suffix, depending on the procedure involved. See the `summary_proceeding_suffixes` table for more information about what each suffix means. When multiple cases are joined together, the proceeding ID will use the case ID of the case that was lodged first (e.g., the case ID with the earliest year and lowest number). In stand-alone cases, the proceeding ID is the same as the case ID, with the possible addition of a suffix. Secondary proceedings always have a suffix. A main proceeding will have a `P` suffix when the procedure is an appeal (from the French *pourvoi*). There are two exceptions to these rules. Proceedings related to an opinion or ruling are not associated with a particular case and have their own ID scheme.

When a proceeding is an extension of a proceeding that already has a suffix, the second suffix comes after the first, separated by a dash. For example, proceeding `C-265/17 P-DEP` at the Court of Justice involves an application as to costs (a dispute between parties about who should pay the costs of litigation) related to an appeal. This proceeding is an extension of proceeding `C-265/17 P`, which is an appeal of General Court case `T-194/13`. When there are multiple, separate procedures of the same type, the Court uses Roman numerals to distinguish between them. For example, there could be a `DEP II`, which would be a second application as to costs, or a `REV II`, which would be a second application for revision.

The review procedure works somewhat differently than other procedures. In the review procedure, the Court reviews a decision made by the General Court or Civil Service Tribunal. Proceeding IDs work slightly differently for this procedure. The suffix `RX` indicates the proceeding in which the Court of Justice issues a decision about whether it will review the decision of the lower court. If the Court decides to review the decision, the suffix `RX II`

indicates the proceeding in which the Court of Justice issues a judgment reviewing that decision. The Court of Justice can refer the case back to the lower court for a ruling on the merits. When this happens, the original proceeding ID is used, but a `RENV-RX` suffix is added. This differentiates a referral back after review from a referral back after an appeal, which would be indicated by `RENV`. If the original proceeding was an appeal, the new suffix would be `P-RENV-RX`.

For example, `F-142/11` was a case before the Civil Service Tribunal. The decision of the Civil Service Tribunal was appealed to the General Court. The proceeding ID of the appeal was `T-130/14 P`. In its judgment, the General Court referred the case back to the Civil Service Tribunal, and that proceeding was `F-142/11 RENV`. The second decision of the Civil Service Tribunal was also appealed to the General Court. The proceeding ID of this second appeal was `T-646/16 P`. The General Court issued a second decision on appeal. Then, the Court of Justice decided to review the second decision of the General Court in `T-646/16 P`. The proceeding ID for the proceeding where the Court of Justice decided whether to review the decision was `C-542/18 RX`. The proceeding ID for the proceeding in which the Court of Justice issued a judgment reviewing the decision for the General Court was `C-542/18 RX II`. The Court of Justice decided to refer the case back to the General Court, and the proceeding ID for that proceeding was `T-646/16 P-RENV-RX`.

Cases

All cases at the CJ, GC, and CST

Description

The `cases` table includes data on the universe of CJEU cases. There is one observation per case. The table includes every case that should exist, based on the highest case number that occurs in each year in the Registry. The `cases` table indicates all proceedings associated with the case, whether the main proceeding in the case is pending (as of 31 December, 2022), is removed from the Registry, is transferred to another court, or is joined to another proceeding, whether is an appeal associated with the main proceeding, whether the main proceeding is an appeal, all decisions that the Court has issued with respect to the case, whether there was a judgment in the case, and whether there was an AG opinion in the case.

The Registry does not contain any records for a small number of cases. However, the `cases` table may still have data on some of these cases because decisions related to them appear in the `decisions` table, which also incorporates data from InfoCuria and EUR-Lex. There are also some cases that should exist, but that do not appear in the Registry, InfoCuria, or EUR-Lex.

Each observation in the `cases` table is uniquely identified by a IUROPA case ID (`iuropa_case_id`) and by a CJEU case ID (`cjeu_case_id`). Note that these ID numbers ignore proceeding suffixes.

IUROPA case IDs have the format `CJEU:X:####:####`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, and the second number is the number of the case (with preceding

zeros). The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

CJEU case IDs have the format `X-####/##`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, and the second number is the last two digits of the year. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it.

We code a set of variables that capture events related to each case. When there are multiple proceedings associated with a case (e.g., a main proceeding and an application as to costs), we code these variables based on the main proceeding.

For each case, we code whether the main proceeding is pending as of 31 December, 2022 (`is_pending`), is removed from the Registry (`is_removed`), is transferred to another Court (`is_transferred`), or is joined to another proceeding (`is_joined`). For cases in which the main proceeding is joined to another proceeding, we code the CJEU case ID of the new proceeding of record (`joined_to`). For cases that are transferred, we code the CJEU case ID of the new case (`transferred_to`), as the case ID will change.

We also code data about appeals. For each case, we code a list of any cases that are appeals related to the case (`list_on_appeal`) and a count of those cases (`count_on_appeal`). For appeals, we code a list of all cases that are being appealed (`list_appeal_of`) and a count of those cases (`count_appeal_of`).

Finally, we code a list of all decisions associated with each case (`list_decisions`), according to the `decisions` table and the count of those decisions (`count_decisions`). Based on this list, we code whether there is at least one judgment associated with the case (`has_judgment`) and whether there is at least one AG opinion associated with the case (`has_opinion`).

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

court_id integer An ID number that uniquely identifies each court. Coded **1** for the Court of Justice, **2** for the General Court, and **3** for the Civil Service Tribunal.

court string The court at which the case was lodged. The possible values are: Court of Justice, General Court, and Civil Service Tribunal.

iuroopa_case_id string The IUROPA case ID number. Uniquely identifies each case. Coded based on **cjeu_case_id**. IUROPA case IDs have the format **CJEU:X:####:####**, where **X** is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, and the second number is the number of the case (with preceding zeros). The possible values of **X** are **C** for the Court of Justice, **T** for the General Court, and **F** for the Civil Service Tribunal.

cjeu_case_id string The CJEU case ID number. Uniquely identifies each case. CJEU case IDs have the format **X-####/##**, where **X** is a letter that indicates the court at which the case was lodged, the first number is the number of the case, and the second number is the last two digits of the year. The possible values of **X** are **C** for the Court of Justice, **T** for the General Court, and **F** for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a **C-** prefix, but the database always includes it.

case_year integer The year that the case was lodged. Coded based on **cjeu_case_id**.

case_number integer The number of the case. Case numbers restart at **1** at the beginning of each year. Coded based on **cjeu_case_id**.

case_name string The name of the case, according to the Registry. Coded **not available** when the Registry does not have a record for the case. If a case has multiple proceedings, the name of the case is the name of the main proceeding.

is_pending dummy Whether the main proceeding in the case is pending, according to the Registry, as of 31 December, 2022.

`is_removed` `dummy` Whether the main proceeding in the case has been removed from the Registry.

`is_transferred` `dummy` Whether the main proceeding in the case has been transferred to another court.

`transferred_to` `string` If the main proceeding in the case has been transferred to another court (see `is_transferred`), the `cjeu_case_id` for the case after the transfer. Coded `none` for cases that have not been transferred. Note that this variable does not include proceeding suffixes.

`is_joined` `dummy` Whether the main proceeding in the case has been joined to another case.

`joined_to` `string` If the main proceeding in the case has been joined to another case (see `is_joined`), the `cjeu_case_id` for the case of record, which is the `cjeu_case_id` of the earliest case lodged. Coded `none` for cases in which the main proceeding has not been joined. Note that this variable does not include proceeding suffixes.

`count_on_appeal` `integer` The number of appeals related to each case. Coded based on `list_on_appeal`.

`list_on_appeal` `string` A list of the appeals related to each case, separated by commas. Coded `not applicable` for all cases at the Court of Justice and coded `none` for cases at the General Court and at the Civil Service Tribunal that are not appealed. Note that this variable does not include proceeding suffixes.

`count_appeal_of` `integer` If the main proceeding in the case is an appeal, the number of case that are being appealed. Coded `0` if the proceeding is not an appeal. Coded based on `list_appeal_of`.

`list_appeal_of` `string` If the case is an appeal, a list of the cases that are being appealed, separated by commas. Coded `not_applicable` for cases that are not appeals. Note that this variable does not include proceeding suffixes.

`count_proceedings` `integer` The number of proceedings associated with this case, according to the `proceedings` table. Coded based on `list_proceedings`.

`list_proceedings` `string` A list of the proceedings associated with this case, separated by commas, according to the `proceedings` table. This variable uses `cjeu_proceeding_id`, which uniquely identifies proceedings.

`count_decisions` `integer` The number of decisions that the Court issued with respect to the case, according to the `decisions` table. Coded based on `list_decisions`.

`list_decisions` `string` A list of the decisions that the Court issued with respect to the case, separated by commas, according to the `decisions` table. Coded `none` if there are no decisions. This variable uses `iuropa_decision_id`, which uniquely identifies decisions.

`has_judgment` `dummy` Whether there was a judgment related to the case. Coded based on `list_decisions`.

`has_opinion` `dummy` Whether there was an AG opinion related to the case. Coded based on `list_decisions`.

Proceedings

All proceedings at the CJ, GC, and CST

Description

The `proceedings` table includes data on the universe of CJEU proceedings. There is one observation per proceeding. A proceeding involves a specific set of cases and procedures. The `cases` table indicates whether the proceeding is pending (as of 31 December, 2022), is removed from the Registry, is transferred to another court, or is joined to another proceeding, whether there is an appeal associated with the proceeding, whether the proceeding is an appeal, all proceedings that have been joined to the proceeding, all decisions that the Court has issued with respect to the proceeding, whether there was a judgment in the proceeding, and whether there was an AG opinion in the proceeding.

Each observation in the `proceedings` table is uniquely identified by a IUROPA proceeding ID (`iuropa_proceeding_id`) and by a CJEU proceeding ID (`cjeu_proceeding_id`).

CJEU proceeding IDs and IUROPA proceeding IDs have a similar format as CJEU case IDs and IUROPA case IDs. IUROPA proceeding IDs have the format `CJEU:X:####:####[Y]`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, the second number is the number of the case (with preceding zeros), and `[Y]` is an optional suffix that distinguishes between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the suffix.

CJEU proceedings IDs have the format `X-####/## Y`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, the second number is the last two digits of the year, and `Y` is an optional suffix that differentiates between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it. Opinions of the Court and rulings use the prefixes `OPINION` and `RULING` instead of `C-`. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`.

Note that the Court is occasionally inconsistent about how it formats suffixes, especially for older proceedings. We standardize these suffixes to make them easier to understand. If you notice a difference between a proceeding ID in the CJEU Database Platform and a proceeding ID in InfoCuria or EUR-Lex, it is because we have standardized it. We provide a supplemental table, `summary_proceeding_suffixes`, that indicates what each suffix means.

We code a set of variables that capture events related to each proceeding. The Registry records when a proceeding is removed from the Registry (which requires an order), when a proceeding is joined to another proceeding, when a proceeding is transferred to another court, and when the Court publishes a decision. However, the Registry does not always include a record for every time a proceeding is joined. We use our `decisions` table, which indicates which cases each decision applies to, to supplement the information on joins in the Registry.

For each proceeding, we code whether the proceeding is pending as of 31 December, 2022 (`is_pending`), is removed from the Registry (`is_removed`), is joined to another proceeding (`is_joined`), or is transferred to another Court (`is_transferred`). For proceedings that are joined, we code the CJEU proceeding ID of the new proceeding of record (`joined_to`). For proceedings that are transferred, we code the CJEU proceeding ID of the new case (`transferred_to`), as the proceeding ID will change.

We also code data about appeals. For each proceeding, we code a list of any proceedings that are appeals related to the proceedings (`list_on_appeal`) and a count of those proceedings (`count_on_appeal`). For appeals, we code a list of all proceedings that are

being appealed (`list_appeal_of`) and a count of those proceedings (`count_appeal_of`).

Finally, we code a list of all decisions associated with each proceeding (`list_decisions`), according to the `decisions` table and the count of those decisions (`count_decisions`). Based on this list, we code whether there is at least one judgment associated with the proceeding (`has_judgment`) and whether there is at least one AG opinion associated with the proceeding (`has_opinion`).

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`court_id` `integer` An ID number that uniquely identifies each court. Coded `1` for the Court of Justice, `2` for the General Court, and `3` for the Civil Service Tribunal.

`court` `string` The court at which the case was lodged. The possible values are: `Court of Justice`, `General Court`, and `Civil Service Tribunal`.

`iuropa_case_id` `string` The IUROPA case ID number. Uniquely identifies each case. Coded based on `cjeu_case_id`. IUROPA case IDs have the format `CJEU:X:####:####`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, and the second number is the number of the case (with preceding zeros). The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`cjeu_case_id` `string` The CJEU case ID number. Uniquely identifies each case. CJEU case IDs have the format `X-####/##`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, and the second number is the last two digits of the year. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it.

`case_year` integer The year that the case was lodged. Coded based on `cjeu_case_id`.

`case_number` integer The number of the case. Case numbers restart at 1 at the beginning of each year. Coded based on `cjeu_case_id`.

`iuropa_proceeding_id` string The IUROPA proceeding ID number. Uniquely identifies each proceeding. Coded based on `cjeu_proceeding_id`. IUROPA proceeding IDs have the format `CJEU:X:####:####[Y]`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, the second number is the number of the case (with preceding zeros), and `[Y]` is an optional suffix that distinguishes between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the suffix.

`cjeu_proceeding_id` string The CJEU proceeding ID number. Uniquely identifies each proceeding. CJEU proceedings IDs have the format `X-####/## Y`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, the second number is the last two digits of the year, and `Y` is an optional suffix that differentiates between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it. Opinions of the Court and rulings use the prefixes `OPINION` and `RULING` instead of `C-`. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`.

`proceeding_suffix` string The proceeding suffix, if applicable. Coded based on `cjeu_proceeding_id`. This suffix indicates the legal procedure(s) associated with the proceeding. Coded `none` if `cjeu_proceeding_id` does not have a suffix. See the

`summary_proceeding_suffixes` table for more information about the possible values and what each suffix means.

`proceeding_name` `string` The name of the proceeding. This is the name of the proceeding of record. If the Court has joined multiple proceedings together, the proceeding of record is the proceeding that was lodged first. The name of the case is presented exactly as it appears in the Registry. Coded `not available` for proceedings that do not have a record in the Registry.

`is_pending` `dummy` Whether the proceeding is pending, according to the Registry, as of 31 December, 2022.

`is_removed` `dummy` Whether the proceeding is removed from the Registry.

`is_transferred` `dummy` Whether the proceeding is transferred to another court.

`transferred_to` `string` If the proceeding is transferred to another court (see `is_transferred`), the `cjeu_proceeding_id` for the proceeding after the transfer. Coded `none` for proceedings that have not been transferred.

`is_joined` `dummy` Whether the proceeding is joined to another proceeding.

`joined_to` `string` If the proceeding is joined to another proceeding (see `is_joined`), the `cjeu_proceeding_id` for the proceeding of record, which is the `cjeu_proceeding_id` of the earliest proceeding lodged. Coded `none` for proceedings that are not joined.

`count_on_appeal` `integer` The number of appeals related to each proceeding. Coded based on `list_on_appeal`.

`list_on_appeal` `string` A list of the appeals related to each proceeding, separated by commas. Coded `not applicable` for all proceedings at the Court of

Justice and coded `none` for proceedings at the General Court and at the Civil Service Tribunal that are not appealed.

`count_appeal_of` `integer` If the proceeding is an appeal, the number of proceedings that are being appealed. Coded `0` if the proceeding is not an appeal. Coded based on `list_appeal_of`.

`list_appeal_of` `string` If the proceeding is an appeal, a list of the proceedings that are being appealed, separated by commas. Coded `not applicable` for proceedings that are not appeals.

`count_joined_proceedings` `integer` The number of proceedings that the Court has joined together. Coded based on `list_joined_proceedings`.

`list_joined_proceedings` `string` A list of the proceedings that the Court has joined together, including the proceeding of record, separated by commas.

`count_decisions` `integer` The number of decisions that the Court issued with respect to the proceeding, according to the `decisions` table. Coded based on `list_decisions`.

`list_decisions` `string` A list of the decisions that the Court issued with respect to the proceeding, separated by commas, according to the `decisions` table. Coded `none` if there are no decisions. This variable uses `iuropa_decision_id`, which uniquely identifies decisions.

`has_judgment` `dummy` Whether there was a judgment related to the proceeding. Coded based on `list_decisions`.

`has_opinion` `dummy` Whether there was an AG opinion related to the proceeding. Coded based on `list_decisions`.

Decisions

All decisions issued by the CJ, GC, and CST

Description

The `decisions` table includes data on the universe of CJEU decisions, including judgments, orders, Advocate-General opinions, and more. There is one observation per decision. The `decisions` table has one observation per CJEU decision. The `decisions` table indicates the ECLI and CELEX numbers for each decision, the type of the decision, and every proceeding that is associated with the decision.

Each observation in the `decisions` table is uniquely identified by a IUROPA decision ID (`iuropa_decision_id`) and by a CJEU decision ID (`cjeu_decision_id`).

The Registry and InfoCuria use ECLI numbers to identify decisions. EUR-Lex uses CELEX numbers and ECLI numbers. However, ECLI numbers do not actually uniquely identify decisions (1) because the Court sometimes incorrectly uses one ECLI number for two distinct decisions and (2) because the Court sometimes assigns multiple ECLI numbers to one unique decision.

CJEU decision IDs are a combination of the ECLI number (`ecli`) and the date of the decision (`decision_date`). CJEU decision IDs have the format `ECLI:EU:X:####:####:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year the decision was published, the second number is the document number assigned by the Court, and the third number is the date in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

IUROPA decision IDs (`iuropa_decision_id`) are a useful alternative to CJEU decision IDs. We provide this ID number (1) because ECLI numbers are not actually uniquely identifying, (2) because ECLI numbers are based on a counter that cannot be recreated from metadata about a decision, and (3) because ECLI numbers do not indicate the type of the decision or the case of record that the decision applies to.

IUROPA decision IDs are an extension of IUROPA case IDs. IUROPA decision IDs have the format `CJEU:X:####:####:Y:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year of the case, the second number is the case number (with preceding zeros), `Y` is a letter that indicates the type of the decision, and the third number is the date of the document in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values of `Y` are based on CELEX descriptors: `J` for judgments, `O` for orders, `V` for opinions of the Court, `D` for decisions, `X` for rulings, `S` for seizure orders, `T` for third-party proceedings, `C` for AG opinions, and `P` for AG views. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the date.

Sometimes the Court issues two decisions of the same type on the same day in the same case. When this occurs, we append the document number from the ECLI number, in brackets, to the end of the IUROPA decision ID. For example, there are two distinct General Court orders in case `T-57/11` from `2012-11-06`, with ECLI numbers `ECLI:EU:T:2012:580` and `ECLI:EU:T:2012:582`, and we distinguish between them using the following IUROPA decision IDs: `CJEU:T:2011:0057:0:20121106[580]` and `CJEU:T:2011:0057:0:20121106[582]`.

Note that there are some orders in the Registry that are not included in the `decisions` table because the Registry does not report the date that the order was published, which means we cannot construct a IUROPA decision ID or a CJEU decision ID for the order. These orders do not appear in InfoCuria or EUR-Lex, and there is no other record of them.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

court_id integer An ID number that uniquely identifies each court. Coded **1** for the Court of Justice, **2** for the General Court, and **3** for the Civil Service Tribunal.

court string The court at which the case was lodged. The possible values are: Court of Justice, General Court, and Civil Service Tribunal.

iuropa_case_id string The IUROPA case ID number. Uniquely identifies each case. Coded based on **cjeu_case_id**. IUROPA case IDs have the format **CJEU:X:####:####**, where **X** is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, and the second number is the number of the case (with preceding zeros). The possible values of **X** are **C** for the Court of Justice, **T** for the General Court, and **F** for the Civil Service Tribunal.

cjeu_case_id string The CJEU case ID number. Uniquely identifies each case. CJEU case IDs have the format **X-####/##**, where **X** is a letter that indicates the court at which the case was lodged, the first number is the number of the case, and the second number is the last two digits of the year. The possible values of **X** are **C** for the Court of Justice, **T** for the General Court, and **F** for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a **C-** prefix, but the database always includes it.

case_year integer The year that the case was lodged. Coded based on **cjeu_case_id**.

case_number integer The number of the case. Case numbers restart at **1** at the beginning of each year. Coded based on **cjeu_case_id**.

iuropa_proceeding_id string The IUROPA proceeding ID number. Uniquely identifies each proceeding. Coded based on **cjeu_proceeding_id**. IUROPA proceeding IDs have the format **CJEU:X:####:####[Y]**, where **X** is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, the second number is the number of the case (with preceding zeros), and **[Y]** is an optional suffix that distinguishes between proceedings associated with the same case. The

possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the suffix.

`cjeu_proceeding_id` `string` The CJEU proceeding ID number. Uniquely identifies each proceeding. CJEU proceedings IDs have the format `X-####/## Y`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, the second number is the last two digits of the year, and `Y` is an optional suffix that differentiates between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it. Opinions of the Court and rulings use the prefixes `OPINION` and `RULING` instead of `C-`. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`.

`proceeding_suffix` `string` The proceeding suffix, if applicable. Coded based on `cjeu_proceeding_id`. This suffix indicates the legal procedure(s) associated with the proceeding. Coded `none` if `cjeu_proceeding_id` does not have a suffix. See the `summary_proceeding_suffixes` table for more information about the possible values and what each suffix means.

`proceeding_name` `string` The name of the proceeding. This is the name of the proceeding of record. If the Court has joined multiple proceedings together, the proceeding of record is the proceeding that was lodged first. The name of the case is presented exactly as it appears in the Registry. Coded `not available` for proceedings that do not have a record in the Registry.

`iuroipa_decision_id` `string` The IUROPA decision ID number. Uniquely identifies each decision. IUROPA decision IDs have the format `CJEU:X:####:####:Y:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year of the case, the second number is the case number (with preceding zeros), `Y` is a

letter that indicates the type of the decision, and the third number is the date of the document in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values of `Y` are based on CELEX descriptors: `J` for judgments, `O` for orders, `V` for opinions of the Court, `D` for decisions, `X` for rulings, `S` for seizure orders, `T` for third-party proceedings, `C` for AG opinions, and `P` for AG views. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the date.

`cjeu_decision_id` `string` The CJEU decision ID number. Uniquely identifies each decision. CJEU decision IDs are a combination of the ECLI number (`ecli`) and the date of the decision (`decision_date`). CJEU decision IDs have the format `ECLI:EU:X:####:####:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year the decision was published, the second number is the document number assigned by the Court, and the third number is the date in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`ecli` `string` The ECLI number for the decision. Does not always uniquely identify each decision. ECLI numbers have the format `ECLI:EU:X:####:####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year the decision was published, and the second number is a document number assigned by the Court. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`celex` `string` The CELEX number for the decision. Coded `not applicable` for decisions that do not appear in EUR-Lex. Does not always uniquely identify each decision. CELEX numbers for CJEU decisions have the format `6####XY####`, where the first number is the year that the decision was published, `XY` is the CELEX descriptor for the decision, and the second number is the number of the case (with preceding zeros). The letter `X` indicates the court and the letter `Y` indicates the type of the decision. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values for `Y` are `J` for judgments, `O` for orders, `V` for opinions of the Court, `D` for decisions, `X` for rulings, `S` for seizure

orders, **T** for third-party proceedings, **C** for AG opinions, and **P** for AG views. Some CELEX numbers also have a suffix with the format **(##)**, where the number is a counter (with proceeding zeros) that differentiates between decisions of the same type related to the same case.

decision_type **string** The type of the decision. The possible values are (in decreasing order of frequency): **judgment**, **order**, **AG opinion**, **AG view**, **opinion**, **decision**, **seizure order**, **third-party proceedings**, and **ruling**. Note that not all types of decisions are possible at all courts. At the Court of Justice, all values are possible. At the General Court, the possible values are: **judgment**, **order**, **third-party proceedings** (rare), and **AG opinion** (rare). At the Civil Service Tribunal, the possible values are **judgment** and **order**.

decision_date **date** The date that the decision was published by the Court in the format **YYYY-MM-DD**.

count_joined_proceedings **integer** The number of proceedings that the Court has joined together. Coded based on **list_joined_proceedings**.

list_joined_proceedings **string** A list of the proceedings that the Court has joined together, including the proceeding of record, separated by commas.

Judges

All judges, AGs, and registrars at the CJ, GC, and CST

Description

The `judges` table includes data on each judge, Advocate-General, and registrar who has served on CJEU. There is one observation per individual. The `judges` table indicates each individual's member state, gender, and professional experience, including whether each judge has prior experience as a judge, academic, civil servant, practicing lawyer, or politician.

Each observation in the `judges` table is uniquely identified by a IUROPA judge ID (`iuropa_judge_id`). IUROPA judge IDs have the format `J:X:##`, where `X` is a code that uniquely identifies the member state of the individual, based on `iuropa_actor_id` from the `summary_actors` table, and the two digits are a counter that uniquely identifies each individual within a member state. The counter is assigned with individuals sorted within member state by the date they first joined the Court (in any position as a judge, AG, or registrar), then by last name, and then by first name. This ensures that new individuals can be added without affecting existing values.

The data for the `judges` table comes from the official biographies of the judges, AGs, and registrars on the CJEU website. We hand code the data for each judge and cross-verify the data with the `appointments`, `assignments`, `judgments`, and `opinions` tables to ensure that every individual in the `judges` table appears in the `appointments` table and that every individual in the `assignments`, `judgements`, and `opinions` tables appears in the `judges` table.

We code the prior professional experience of each judge using 5 dummy variables:

`was_judge`, `was_academic`, `was_civil_servant`, `was_lawyer`, and `was_politician`. We apply the following coding rules:

- We only consider prior experience indicated in an individual's official biography on the CJEU website.
- We only consider professional experience before an individual's first appointment to the CJEU in any position (judge, AG, or registrar).
- We code `was_judge` as 1 for individuals who have been a judge (including magistrate judges) at the international, national, or sub-national level. We do not consider prior experience at the CJEU in coding this variable.
- We code `was_academic` as 1 for individuals who have held a permanent academic position. We do not consider temporary or visiting positions in coding this variable.
- We code `was_civil_servant` as 1 for individuals who have worked for domestic or international governmental institutions in non-elected positions. This includes positions at any rank, from government lawyers to high-level government ministers. We do not count judges as civil servants.
- We code `was_lawyer` as 1 for individuals who have worked as a lawyer at a law firm, a company, a court, or a governmental institution. Individuals who have worked as law clerk (or similar) at a court or as legal counsel at a government institution are also coded as civil servants.
- We code `was_politician` as 1 for individuals who have held an elected position in their member state (at the national level or at the sub-national level) or who have been members of the European Parliament.

Variables

`key_id` integer An ID number that uniquely identifies each observation.

`iuropa_judge_id` `string` The IUROPA judge ID number. Uniquely identifies each judge, AG, and registrar. IUROPA judge IDs have the format `J:X:##`, where `X` is a code that uniquely identifies the member state of the individual, based on `iuropa_actor_id` from the `summary_actors` table, and the number is a counter that uniquely identifies each individual within a member state. The counter is assigned with individuals sorted within member state by the date they first joined the Court (in any position as a judge, AG, or registrar), then by last name, and then by first name.

`full_name` `string` The full name of the individual (including any diacritical marks).

`first_name` `string` The first name of the individual (including any diacritical marks).

`last_name` `string` The last name of the individual (including any diacritical marks).

`last_name_latin` `string` The last name of the individual using only Latin (ASCII) characters.

`last_name_label` `string` A label that uniquely identifies each individual, by including the individual's first initial for individuals with the same last name.

`last_name_latin_label` `string` A label that uniquely identifies each individual, by including the individual's first initial for individuals with the same last name, using only Latin (ASCII) characters.

`iuropa_member_state_id` `string` The IUROPA actor ID number for the individual's member state. Coded based on `iuropa_actor_id` in the `summary_actors` table.

`member_state` `string` The member state of the individual.

`birth_year` `integer` The birth year of the individual.

`is_female` `dummy` Whether the individual is female.

`is_current` `dummy` Whether the individual is a current member of the CJEU as a judge, Advocate General (AG), or registrar.

`start_date` `date` The date that the individual started their first position at the CJEU in the format `YYYY-MM-DD`.

`end_date` `date` The date that the individual ended their last position at the CJEU in the format `YYYY-MM-DD`. Coded `2022-12-31` if the judge is a current member of the Court.

`was_judge` `dummy` Whether the judge had professional experience as a judge before joining the Court.

`was_academic` `dummy` Whether the judge had professional experience as an academic before joining the Court.

`was_civil_servant` `dummy` Whether the judge had professional experience as a civil servant before joining the Court.

`was_lawyer` `dummy` Whether the judge had professional experience as a practicing lawyer before joining the Court.

`was_politician` `dummy` Whether the judge had professional experience as a politician before joining the Court.

Judgments

All judgments issued by the CJ, GC, and CST

Description

The `judgments` table includes data on the universe of CJEU judgments. There is one observation per judgment. The `judgments` table is a sub-table of the `decisions` table that includes variables that are specific to judgments. The `judgments` table indicates the legal procedure(s), the formation of the Court, the judges on the panel, the Judge-Rapporteur, and the Advocate General (AG), whether the Court requested an AG opinion, the documents cited in the judgment, the subject matter of the judgment, and more.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`court_id` `integer` An ID number that uniquely identifies each court. Coded `1` for the Court of Justice, `2` for the General Court, and `3` for the Civil Service Tribunal.

`court` `string` The court at which the case was lodged. The possible values are: `Court of Justice`, `General Court`, and `Civil Service Tribunal`.

`iuropa_case_id` `string` The IUROPA case ID number. Uniquely identifies each case. Coded based on `cjeu_case_id`. IUROPA case IDs have the format `CJEU:X:####:####`, where `X` is a letter that indicates the court at which the case was

lodged, the first number is the year that the decision was published, and the second number is the number of the case (with preceding zeros). The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`cjeu_case_id` `string` The CJEU case ID number. Uniquely identifies each case. CJEU case IDs have the format `X-####/##`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, and the second number is the last two digits of the year. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it.

`case_year` `integer` The year that the case was lodged. Coded based on `cjeu_case_id`.

`case_number` `integer` The number of the case. Case numbers restart at `1` at the beginning of each year. Coded based on `cjeu_case_id`.

`iuropa_proceeding_id` `string` The IUROPA proceeding ID number. Uniquely identifies each proceeding. Coded based on `cjeu_proceeding_id`. IUROPA proceeding IDs have the format `CJEU:X:####:####[Y]`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, the second number is the number of the case (with preceding zeros), and `[Y]` is an optional suffix that distinguishes between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the suffix.

`cjeu_proceeding_id` `string` The CJEU proceeding ID number. Uniquely identifies each proceeding. CJEU proceedings IDs have the format `X-####/## Y`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number

of the case, the second number is the last two digits of the year, and `Y` is an optional suffix that differentiates between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it. Opinions of the Court and rulings use the prefixes `OPINION` and `RULING` instead of `C-`. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`.

`proceeding_suffix` `string` The proceeding suffix, if applicable. Coded based on `cjeu_proceeding_id`. This suffix indicates the legal procedure(s) associated with the proceeding. Coded `none` if `cjeu_proceeding_id` does not have a suffix. See the `summary_proceeding_suffixes` table for more information about the possible values and what each suffix means.

`proceeding_name` `string` The name of the proceeding. This is the name of the proceeding of record. If the Court has joined multiple proceedings together, the proceeding of record is the proceeding that was lodged first. The name of the case is presented exactly as it appears in the Registry. Coded `not available` for proceedings that do not have a record in the Registry.

`iuropa_decision_id` `string` The IUROPA decision ID number. Uniquely identifies each decision. IUROPA decision IDs have the format `CJEU:X:####:####:Y:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year of the case, the second number is the case number (with preceding zeros), `Y` is a letter that indicates the type of the decision, and the third number is the date of the document in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values of `Y` are based on CELEX descriptors: `J` for judgments, `O` for orders, `V` for opinions of the Court, `D` for decisions, `X` for rulings, `S` for seizure orders, `T` for third-party proceedings, `C` for AG opinions, and `P` for AG views. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the date.

cjeu_decision_id string The CJEU decision ID number. Uniquely identifies each decision. CJEU decision IDs are a combination of the ECLI number (**ecli**) and the date of the decision (**decision_date**). CJEU decision IDs have the format **ECLI:EU:X:####:####:#####**, where **X** is a letter that indicates the court that issued the decision, the first number is the year the decision was published, the second number is the document number assigned by the Court, and the third number is the date in the format **YYYYMMDD**. The possible values of **X** are **C** for the Court of Justice, **T** for the General Court, and **F** for the Civil Service Tribunal.

ecli string The ECLI number for the decision. Does not always uniquely identify each decision. ECLI numbers have the format **ECLI:EU:X:####:####**, where **X** is a letter that indicates the court that issued the decision, the first number is the year the decision was published, and the second number is a document number assigned by the Court. The possible values of **X** are **C** for the Court of Justice, **T** for the General Court, and **F** for the Civil Service Tribunal.

celex string The CELEX number for the decision. Coded **not applicable** for decisions that do not appear in EUR-Lex. Does not always uniquely identify each decision. CELEX numbers for CJEU decisions have the format **6####XY####**, where the first number is the year that the decision was published, **XY** is the CELEX descriptor for the decision, and the second number is the number of the case (with preceding zeros). The letter **X** indicates the court and the letter **Y** indicates the type of the decision. The possible values of **X** are **C** for the Court of Justice, **T** for the General Court, and **F** for the Civil Service Tribunal. The possible values for **Y** are **J** for judgments, **O** for orders, **V** for opinions of the Court, **D** for decisions, **X** for rulings, **S** for seizure orders, **T** for third-party proceedings, **C** for AG opinions, and **P** for AG views. Some CELEX numbers also have a suffix with the format **(##)**, where the number is a counter (with preceding zeros) that differentiates between decisions of the same type related to the same case.

decision_type string The type of the decision. The possible values are (in decreasing order of frequency): **judgment**, **order**, **AG opinion**, **AG view**, **opinion**, **decision**, **seizure order**, **third-party proceedings**, and **ruling**. Note that not all types of decisions are possible at all courts. At the Court of

Justice, all values are possible. At the General Court, the possible values are: `judgment`, `order`, `third-party proceedings` (rare), and `AG opinion` (rare). At the Civil Service Tribunal, the possible values are `judgment` and `order`.

`decision_date` `date` The date that the decision was published by the Court in the format `YYYY-MM-DD`.

`proceeding_date` `date` The date that the proceeding of record was lodged at the Court in the format `YYYY-MM-DD`.

`duration_days` `numeric` The duration of the proceeding in days. The difference between `proceeding_date` and `decision_date`.

`duration_weeks` `numeric` The duration of the proceeding in weeks. The difference between `proceeding_date` and `decision_date`.

`duration_months` `numeric` The duration of the proceeding in months. The difference between `proceeding_date` and `decision_date`.

`duration_years` `numeric` The duration of the proceeding in years. The difference between `proceeding_date` and `decision_date`.

`count_joined_proceedings` `integer` The number of proceedings that the judgment applies to, including the proceeding of record. Coded based on `list_joined_proceedings`.

`list_joined_proceedings` `string` A list of the proceedings that the judgment applies to, separated by commas. Coded based on `cjeu_proceeding_id` from the `proceedings` table.

`count_on_appeal` `integer` The number of appeals related to each proceeding. Coded based on `list_on_appeal`.

`list_on_appeal` `string` A list of the appeals related to each proceeding, separated by commas. Coded `not applicable` for all proceedings at the Court of Justice and coded `none` for proceedings at the General Court and at the Civil Service Tribunal that are not appealed.

`count_appeal_of` `integer` If the proceeding is an appeal, the number of proceedings that are being appealed. Coded `0` if the proceeding is not an appeal. Coded based on `list_appeal_of`.

`list_appeal_of` `string` If the proceeding is an appeal, a list of the proceedings that are being appealed, separated by commas. Coded `not applicable` for proceedings that are not appeals.

`count_authentic_languages` `integer` The number of authentic languages. Coded based on `list_authentic_languages`.

`list_authentic_languages` `string` A list of the authentic languages, separated by commas. Generally, the authentic languages are the languages of the parties.

`count_procedures` `integer` The number of legal procedures associated with the decision. Coded based on `list_procedures`.

`list_procedures` `string` A list of the legal procedures associated with the decision, separated by commas. The outcome is listed in parentheses. The possible legal procedures are (in alphabetical order): `action for annulment`, `action for damages`, `action for failure to act`, `action for failure to fulfill obligations`, `appeal`, `appeal against a penalty`, `application for interim measures`, `application for interpretation`, `application for measures of inquiry`, `application for revision`, `application to intervene`, `application to set aside by default`, `arbitration clause`, `reference for a preliminary ruling`, `responsibility of a member of the institutions`, `review`, and

`staff_case`. The possible outcomes are: `successful`, `unfounded`, `inadmissible`, `unnecessary`, and `interlocutory`. The outcome is omitted for reference for a preliminary ruling and `arbitration clause` because there is not a disposition. Coded based on `procedure` in the `procedures` table.

`is_preliminary_ruling` `dummy` Whether the legal procedure is a reference for a preliminary ruling. Coded based on `list_procedures`.

`is_urgent_procedure` `dummy` Whether the legal procedure is the urgent reference for a preliminary ruling procedure. Coded based on `list_procedures` and `proceeding_suffix`.

`is_direct_action` `dummy` Whether the legal procedure is direct action. Direct actions include actions for annulment, actions for failure to act, actions for failure to fulfill obligations, and actions for damages. Coded based on `list_procedures`.

`is_appeal` `dummy` Whether the legal procedure is an appeal. Coded based on `list_procedures`.

`is_staff_case` `dummy` Whether the legal procedure is a staff case. Coded based on `list_procedures`.

`is_referral_back` `dummy` Whether the proceeding associated with the decision has been referred back after an appeal or review. Coded based on `proceeding_suffix`.

`count_referring_member_states` `integer` When the legal procedure is a reference for a preliminary ruling, the number of member states referring questions to the Court. Coded `0` otherwise. Coded based on `list_referring_member_states`.

`list_referring_member_states` `string` When the legal procedure is a reference for a preliminary ruling, a list of the member states referring questions to the Court, separated by commas. For decisions that apply to multiple proceedings, there can be

multiple member states. Coded `not applicable` when the legal procedure is not a reference for a preliminary ruling.

`count_observers` `integer` The number of institutional observers. Coded `0` when the legal procedure is not a reference for a preliminary ruling. Coded based on `list_observers`.

`list_observers` `string` A list of institutional observers, separated by a comma. Coded `not applicable` when the legal procedure is not a reference for a preliminary ruling. Coded based on `observer` in the `observers` table.

`iuropa_formation_id` `string` The IUROPA formation ID. Uniquely identifies each formation of the Court. IUROPA formation IDs have the format `F:X:Y`, where `X` indicates the Court and `Y` indicates the formation. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values for `Y` are `01` through `10` for numbered chambers, `FC` for Full Court, `GC` for Grand Chamber, `AC` for Appeals Chamber, and `SJ` for Single Judge. Coded `no record` for judgments that are not available in InfoCuria or EUR-Lex.

`formation` `string` The formation of the Court that issued the decision. At the Court of Justice, the possible values are: `First Chamber` through `Tenth Chamber`, `Grand Chamber`, and `Full Court`. At the General Court, the possible values are: `First Chamber` through `Tenth Chamber`, `First Chamber (extended composition)` through `Tenth Chamber (extended composition)`, `Appeals Chamber`, `Grand Chamber`, `Full Court`, and `Single Judge`. At the Civil Service Tribunal, the possible values are: `First Chamber` through `Third Chamber`, `Full Court`, and `Single Judge`. Coded `no record` for judgments that are not available in InfoCuria or EUR-Lex.

`count_judges` `integer` The number of judges on the panel that issued the judgment. There must always be an odd number of judges. Coded based on `list_judges`.

`list_judges` `string` A list of the judges on the panel that issued the judgment in alphabetical order, separated by semi-colons. Calculated based on `judge` in the `assignments` table, which is based on `last_name_label` in the `judges` table.

`iuropa_judge_rapporteur_id` `string` The IUROPA judge ID for the Judge-Rapporteur. Uniquely identifies each judge. Coded `not available` for judgments where the text of the judgment does not indicate who the Judge-Rapporteur is. Coded `no record` for judgments that are not available in InfoCuria or EUR-Lex.

`judge_rapporteur` `string` The name of the Judge-Rapporteur. Coded based on `last_name_label` from the `judges` table. Coded `not available` for judgments where the text of the judgment does not indicate who the Judge-Rapporteur is. Coded `no record` for judgments that are not available in InfoCuria or EUR-Lex.

`count_advocates_general` `integer` The number of AGs assigned to the proceeding. Coded based on `list_advocates_general`.

`list_advocates_general` `string` A list of the AGs assigned to the proceeding, separated by semi-colons. Coded based on `last_name_label` from the `judges` table. In most proceedings, there is only one AGs. Coded `not applicable` for the General Court and the Civil Service Tribunal.

`count_opinions` `integer` The number of AG opinions associated with the proceeding. Coded based on `list_opinions`.

`list_opinions` `string` A list of the AG opinions associated with the proceeding, separated by commas. Coded based on `iuropa_decision_id` from the `decisions` table. Coded `none` if there are not an AG opinion. Coded `not applicable` for the General Court and the Civil Service Tribunal.

`count_cited_documents` `integer` The number of unique documents cited in each judgment. Coded based on `list_cited_documents`.

`list_cited_documents` `string` A list of the unique documents cited in each judgment, separated by commas. Coded based on `cited_celex` from the `citations` table. Coded `not applicable` if EUR-Lex does not provide metadata on citations for the judgment.

`count_subject_keywords` `integer` The number of subject keywords assigned to the judgment in EUR-Lex. Coded based on `list_subject_keywords`.

`list_subject_keywords` `string` The subject keywords assigned to the judgment in EUR-Lex, separated by semi-colons. Coded `not applicable` if EUR-Lex does not provide metadata on subject keywords for the judgment.

`count_directory_codes` `integer` The number of case law directory codes assigned to the judgment in EUR-Lex. Coded based on `list_directory_codes`.

`list_directory_codes` `string` A list of the case law directory codes assigned to the judgment in EUR-Lex, separated by commas. Coded `not applicable` if EUR-Lex does not provide metadata on case law directory codes for the judgment.

Opinions

All AG opinions requested by the CJ

Description

The `opinions` table includes data on the universe of Advocate General (AG) opinions. There is one observation per opinion. The `opinions` table is a sub-table of the `decisions` table that includes decision-level variables that are specific to AG opinions. The `opinions` table indicates the AG who wrote the opinion and all judgments associated with the proceeding, according to the `decisions` table.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`court_id` `integer` An ID number that uniquely identifies each court. Coded `1` for the Court of Justice, `2` for the General Court, and `3` for the Civil Service Tribunal.

`court` `string` The court at which the case was lodged. The possible values are: `Court of Justice`, `General Court`, and `Civil Service Tribunal`.

`iuropa_case_id` `string` The IUROPA case ID number. Uniquely identifies each case. Coded based on `cjeu_case_id`. IUROPA case IDs have the format `CJEU:X:####:####`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, and the second number

is the number of the case (with preceding zeros). The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`cjeu_case_id` `string` The CJEU case ID number. Uniquely identifies each case. CJEU case IDs have the format `X-####/##`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, and the second number is the last two digits of the year. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it.

`case_year` `integer` The year that the case was lodged. Coded based on `cjeu_case_id`.

`case_number` `integer` The number of the case. Case numbers restart at `1` at the beginning of each year. Coded based on `cjeu_case_id`.

`iuropa_proceeding_id` `string` The IUROPA proceeding ID number. Uniquely identifies each proceeding. Coded based on `cjeu_proceeding_id`. IUROPA proceeding IDs have the format `CJEU:X:####:####[Y]`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, the second number is the number of the case (with preceding zeros), and `[Y]` is an optional suffix that distinguishes between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the suffix.

`cjeu_proceeding_id` `string` The CJEU proceeding ID number. Uniquely identifies each proceeding. CJEU proceedings IDs have the format `X-####/## Y`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, the second number is the last two digits of the year, and `Y` is an optional

suffix that differentiates between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it. Opinions of the Court and rulings use the prefixes `OPINION` and `RULING` instead of `C-`. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`.

`proceeding_suffix` `string` The proceeding suffix, if applicable. Coded based on `cjeu_proceeding_id`. This suffix indicates the legal procedure(s) associated with the proceeding. Coded `none` if `cjeu_proceeding_id` does not have a suffix. See the `summary_proceeding_suffixes` table for more information about the possible values and what each suffix means.

`proceeding_name` `string` The name of the proceeding. This is the name of the proceeding of record. If the Court has joined multiple proceedings together, the proceeding of record is the proceeding that was lodged first. The name of the case is presented exactly as it appears in the Registry. Coded `not available` for proceedings that do not have a record in the Registry.

`iuropa_decision_id` `string` The IUROPA decision ID number. Uniquely identifies each decision. IUROPA decision IDs have the format `CJEU:X:####:####:Y:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year of the case, the second number is the case number (with preceding zeros), `Y` is a letter that indicates the type of the decision, and the third number is the date of the document in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values of `Y` are based on CELEX descriptors: `J` for judgments, `O` for orders, `V` for opinions of the Court, `D` for decisions, `X` for rulings, `S` for seizure orders, `T` for third-party proceedings, `C` for AG opinions, and `P` for AG views. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the date.

`cjeu_decision_id` `string` The CJEU decision ID number. Uniquely identifies each decision. CJEU decision IDs are a combination of the ECLI number (`ecli`) and the date of the decision (`decision_date`). CJEU decision IDs have the format `ECLI:EU:X:####:####:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year the decision was published, the second number is the document number assigned by the Court, and the third number is the date in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`ecli` `string` The ECLI number for the decision. Does not always uniquely identify each decision. ECLI numbers have the format `ECLI:EU:X:####:####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year the decision was published, and the second number is a document number assigned by the Court. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`celex` `string` The CELEX number for the decision. Coded `not applicable` for decisions that do not appear in EUR-Lex. Does not always uniquely identify each decision. CELEX numbers for CJEU decisions have the format `6####XY####`, where the first number is the year that the decision was published, `XY` is the CELEX descriptor for the decision, and the second number is the number of the case (with preceding zeros). The letter `X` indicates the court and the letter `Y` indicates the type of the decision. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values for `Y` are `J` for judgments, `O` for orders, `V` for opinions of the Court, `D` for decisions, `X` for rulings, `S` for seizure orders, `T` for third-party proceedings, `C` for AG opinions, and `P` for AG views. Some CELEX numbers also have a suffix with the format `(##)`, where the number is a counter (with preceding zeros) that differentiates between decisions of the same type related to the same case.

`decision_type` `string` The type of the decision. The possible values are (in decreasing order of frequency): `judgment`, `order`, `AG opinion`, `AG view`, `opinion`, `decision`, `seizure order`, `third-party proceedings`, and `ruling`. Note that not all types of decisions are possible at all courts. At the Court of

Justice, all values are possible. At the General Court, the possible values are: `judgment`, `order`, `third-party proceedings` (rare), and `AG opinion` (rare). At the Civil Service Tribunal, the possible values are `judgment` and `order`.

`decision_date` `date` The date that the decision was published by the Court in the format `YYYY-MM-DD`.

`iuropa_advocate_general_id` `string` An ID number that uniquely identifies each AG. Coded based on `iuropa_judge_id` in the `judges` table.

`advocate_general` `string` The name of the AG. Coded based on `last_name_label` from the `judges` table.

`count_judgments` `integer` The number of unique judgments that the Court issued in the proceeding.

`list_judgments` `string` A list of the judgments that the Court issued in the proceeding, separated by commas. Coded based on `iuropa_decision_id` from the `decisions` table.

Appointments

All appointments of judges, AGs, and registrars to the CJ, GC, and CST

Description

The `appointments` table includes data on the appointments of each judge, advocate-general, and registrar. There is one observation per appointment per individual. The `appointments` table indicates the Court to which the individual has been appointed, the position to which the individual has been appointed, the start and end dates of each appointment, and the duration of each appointment in days, weeks, months, and years.

The data for the `appointments` table comes from the official biographies of the judges, AGs, and registrars on the CJEU website. We hand code the data for each judge and cross-verify the data with the `judges` table to ensure that the `appointments` table includes every individual in the `judges` table.

The `appointments` table does not count re-appointments as separate appointments. For example, there will be one observation for an individual who is appointed as a judge and is then reappointed to the same position. However, if an individual moves to a new position (e.g., if an AG becomes a judge or a judge moves from the General Court to the Court of Justice), a new observation is created. If an individual is appointed to the same position for non-consecutive terms, there will be multiple observations.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`iuropa_judge_id` `string` The IUROPA judge ID number. Uniquely identifies each judge, AG, and registrar. IUROPA judge IDs have the format `J:X:##`, where `X` is a code that uniquely identifies the member state of the individual, based on `iuropa_actor_id` from the `summary_actors` table, and the number is a counter that uniquely identifies each individual within a member state. The counter is assigned with individuals sorted within member state by the date they first joined the Court (in any position as a judge, AG, or registrar), then by last name, and then by first name.

`full_name` `string` The full name of the individual (including any diacritical marks).

`first_name` `string` The first name of the individual (including any diacritical marks).

`last_name` `string` The last name of the individual (including any diacritical marks).

`last_name_latin` `string` The last name of the individual using only Latin (ASCII) characters.

`last_name_label` `string` A label that uniquely identifies each individual, by including the individual's first initial for individuals with the same last name.

`last_name_latin_label` `string` A label that uniquely identifies each individual, by including the individual's first initial for individuals with the same last name, using only Latin (ASCII) characters.

`iuropa_member_state_id` `string` The IUROPA actor ID number for the individual's member state. Coded based on `iuropa_actor_id` in the `summary_actors` table.

`member_state` `string` The member state of the individual.

`birth_year` `integer` The birth year of the individual.

`is_female` `dummy` Whether the individual is female.

court string The court to which the individual has been appointed. The possible values are: Court of Justice, General Court, and Civil Service Tribunal.

position string The position to which the individual has been appointed. The possible values are: judge, Advocate General, registrar, President, and Vice-President.

is_current dummy Whether the individual is a current member of the CJEU as a judge, Advocate General (AG), or registrar.

start_date date The date that the individual started the position in the format YYYY-MM-DD.

end_date date The date that the individual ended the position in the format YYYY-MM-DD. Coded 2022-12-31 if the judge is a current member of the Court.

duration_days numeric The duration of the appointment in days based on start_date and end_date.

duration_weeks numeric The duration of the appointment in weeks based on start_date and end_date, rounded to 2 decimal places.

duration_months numeric The duration of the appointment in months based on start_date and end_date, rounded to 2 decimal places.

duration_years numeric The duration of the appointment in years based on start_date and end_date, rounded to 2 decimal places.

Parties

All parties in CJ, GC, and CST cases

Description

The `parties` table includes data about the parties in each proceeding. There is one observation per party per proceeding that has a record in the Registry. The `parties` table indicates the name of each party, as recorded in the Registry, the role of each party in the case (as applicant, defendant, or litigant), and the type of the party (EU member state, EU institution, EU agency, legal person, etc.).

The data for the `parties` table comes from the Registry. The Registry has records for (nearly) every CJEU decision. It also contains records for joins, removals, and pending proceedings. Every proceeding in the `proceedings` table that has a record in the Registry appears in the `parties` table. The `parties` table has one observation for every party mentioned in the proceeding name in the Registry.

Note that the proceeding names in the Registry do not always indicate all parties. Sometimes, a proceeding name will say “and others,” and the “others” are only listed in the text of the decision. Also, in references for a preliminary ruling, the proceeding name in the Registry often only includes the name of one party (i.e., the name of the party by which the case is commonly known), rather than the names of the applicant and defendant before the national court. When this happens, we code the role of the party as `litigant` in the `parties` table. In direct actions, appeals, and staff cases, the proceedings names in the Registry always indicate an applicant and a defendant. Note also that party names in the Registry are not as detailed as party names in the text of decisions. For example, party

names in the Registry do not usually include the abbreviations that indicate the legal status of a company (e.g., Ltd, plc, SA, BV, etc.).

In a future update, we plan to add a more detailed version of the `parties` table that includes every party mentioned in the text of each decision, not just the parties included in the proceeding name in the Registry. We also plan to re-code the `legal person` category in the `party_type` variable using a more detailed classification scheme that differentiates between national institutions, sub-national institutions, organizations, and companies, among other categories.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`court_id` `integer` An ID number that uniquely identifies each court. Coded `1` for the Court of Justice, `2` for the General Court, and `3` for the Civil Service Tribunal.

`court` `string` The court at which the case was lodged. The possible values are: `Court of Justice`, `General Court`, and `Civil Service Tribunal`.

`iuropa_case_id` `string` The IUROPA case ID number. Uniquely identifies each case. Coded based on `cjeu_case_id`. IUROPA case IDs have the format `CJEU:X:####:####`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, and the second number is the number of the case (with preceding zeros). The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`cjeu_case_id` `string` The CJEU case ID number. Uniquely identifies each case. CJEU case IDs have the format `X-####/##`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, and the second number is the last two digits of the year. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of

Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it.

`case_year` `integer` The year that the case was lodged. Coded based on `cjeu_case_id`.

`case_number` `integer` The number of the case. Case numbers restart at `1` at the beginning of each year. Coded based on `cjeu_case_id`.

`iuropa_proceeding_id` `string` The IUROPA proceeding ID number. Uniquely identifies each proceeding. Coded based on `cjeu_proceeding_id`. IUROPA proceeding IDs have the format `CJEU:X:####:####[Y]`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, the second number is the number of the case (with preceding zeros), and `[Y]` is an optional suffix that distinguishes between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the suffix.

`cjeu_proceeding_id` `string` The CJEU proceeding ID number. Uniquely identifies each proceeding. CJEU proceedings IDs have the format `X-####/## Y`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, the second number is the last two digits of the year, and `Y` is an optional suffix that differentiates between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it. Opinions of the Court and rulings use the prefixes `OPINION` and `RULING` instead of `C-`. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`.

`proceeding_suffix` `string` The proceeding suffix, if applicable. Coded based on `cjeu_proceeding_id`. This suffix indicates the legal procedure(s) associated with the proceeding. Coded `none` if `cjeu_proceeding_id` does not have a suffix. See the `summary_proceeding_suffixes` table for more information about the possible values and what each suffix means.

`proceeding_name` `string` The name of the proceeding. This is the name of the proceeding of record. If the Court has joined multiple proceedings together, the proceeding of record is the proceeding that was lodged first. The name of the case is presented exactly as it appears in the Registry. Coded `not available` for proceedings that do not have a record in the Registry.

`party_role` `string` The role of the party in the proceeding. The possible values are `applicant`, `defendant`, and `litigant`. When `proceeding_name` only indicates one party, that party is coded as `litigant`. This is common for references for a preliminary ruling.

`party` `string` The name of the party. For institutional actors that are included in the `summary_actors` table, the name of the actor has been standardized and may differ from the name in `proceeding_name`.

`party_type` `string` The type of the party. Coded based on `actor_type` in the `summary_actors` table with one additional category: `legal person`. The values that occur in the `parties` table are: `EU (general)`, `EU institution`, `EU body`, `EU executive agency`, `EU decentralized agency`, `EU member state`, `EFTA member state`, and `legal person`.

Procedures

All legal procedures associated with CJ, GC, and CST judgments

Description

The `procedures` table includes data on the legal procedures associated with all CJEU judgments. There is one observation per legal procedure per judgment. There are many legal procedures, but the major ones are references for preliminary rulings, actions for failure to fulfill obligations, actions for annulment, actions for failure to act, damages for non-contractual liability, appeals, and staff cases. The `procedures` table also indicates the ruling of the Court with respect to each legal procedure.

The data for the `procedures` table comes from EUR-Lex metadata. When EUR-Lex metadata is not available, we code the information by hand based on the text of the judgment.

This key variable in the `procedures` table is `procedure`, which indicates each legal procedure associated with a judgment. The `summary_procedures` table indicates all of the possible legal procedures and their legal basis under the Treaties, the Statute of the Court of Justice of the European Union, and the Rules of Procedure.

There are five possible outcomes with respect to each procedure, which we indicate using dummy variables: `is_successful` (the Court has sided with the plaintiff on the merits), `is_unfounded` (the Court has sided with the defendant on the merits), `is_inadmissible` (the Court has ruled that the application filed by the plaintiff is inadmissible), `is_unnecessary` (the Court has ruled that it is not necessary to render a decision), and `is_interlocutory` (the Court has issued a temporary decision and reserves judgment on the application). Note that there can be multiple outcomes with

respect to a legal procedure. For example, an action for annulment could be partially inadmissible and partially unfounded, in which case `is_inadmissible` would be coded `1` and `is_unfounded` would be coded `1`.

There are two legal procedures where the ruling of the Court does not have a disposition: `reference for a preliminary ruling` and `arbitration clause`. For these procedures, the dummy variables that indicate the outcome will all be coded `0` and the dummy variable `no_disposition` will be coded `1`.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`court_id` `integer` An ID number that uniquely identifies each court. Coded `1` for the Court of Justice, `2` for the General Court, and `3` for the Civil Service Tribunal.

`court` `string` The court at which the case was lodged. The possible values are: `Court of Justice`, `General Court`, and `Civil Service Tribunal`.

`iuropa_case_id` `string` The IUROPA case ID number. Uniquely identifies each case. Coded based on `cjeu_case_id`. IUROPA case IDs have the format `CJEU:X:####:####`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, and the second number is the number of the case (with preceding zeros). The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`cjeu_case_id` `string` The CJEU case ID number. Uniquely identifies each case. CJEU case IDs have the format `X-####/##`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, and the second number is the last two digits of the year. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it.

`case_year` integer The year that the case was lodged. Coded based on `cjeu_case_id`.

`case_number` integer The number of the case. Case numbers restart at 1 at the beginning of each year. Coded based on `cjeu_case_id`.

`iuropa_proceeding_id` string The IUROPA proceeding ID number. Uniquely identifies each proceeding. Coded based on `cjeu_proceeding_id`. IUROPA proceeding IDs have the format `CJEU:X:####:####[Y]`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, the second number is the number of the case (with preceding zeros), and `[Y]` is an optional suffix that distinguishes between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the suffix.

`cjeu_proceeding_id` string The CJEU proceeding ID number. Uniquely identifies each proceeding. CJEU proceedings IDs have the format `X-####/## Y`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, the second number is the last two digits of the year, and `Y` is an optional suffix that differentiates between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it. Opinions of the Court and rulings use the prefixes `OPINION` and `RULING` instead of `C-`. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`.

`proceeding_suffix` string The proceeding suffix, if applicable. Coded based on `cjeu_proceeding_id`. This suffix indicates the legal procedure(s) associated with the proceeding. Coded `none` if `cjeu_proceeding_id` does not have a suffix. See the

`summary_proceeding_suffixes` table for more information about the possible values and what each suffix means.

`proceeding_name` `string` The name of the proceeding. This is the name of the proceeding of record. If the Court has joined multiple proceedings together, the proceeding of record is the proceeding that was lodged first. The name of the case is presented exactly as it appears in the Registry. Coded `not available` for proceedings that do not have a record in the Registry.

`iuropa_decision_id` `string` The IUROPA decision ID number. Uniquely identifies each decision. IUROPA decision IDs have the format `CJEU:X:####:####:Y:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year of the case, the second number is the case number (with preceding zeros), `Y` is a letter that indicates the type of the decision, and the third number is the date of the document in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values of `Y` are based on CELEX descriptors: `J` for judgments, `O` for orders, `V` for opinions of the Court, `D` for decisions, `X` for rulings, `S` for seizure orders, `T` for third-party proceedings, `C` for AG opinions, and `P` for AG views. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the date.

`cjeu_decision_id` `string` The CJEU decision ID number. Uniquely identifies each decision. CJEU decision IDs are a combination of the ECLI number (`ecli`) and the date of the decision (`decision_date`). CJEU decision IDs have the format `ECLI:EU:X:####:####:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year the decision was published, the second number is the document number assigned by the Court, and the third number is the date in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`ecli` `string` The ECLI number for the decision. Does not always uniquely identify each decision. ECLI numbers have the format `ECLI:EU:X:####:####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year the

decision was published, and the second number is a document number assigned by the Court. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`celex` `string` The CELEX number for the decision. Coded `not applicable` for decisions that do not appear in EUR-Lex. Does not always uniquely identify each decision. CELEX numbers for CJEU decisions have the format `6####XY####`, where the first number is the year that the decision was published, `XY` is the CELEX descriptor for the decision, and the second number is the number of the case (with preceding zeros). The letter `X` indicates the court and the letter `Y` indicates the type of the decision. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values for `Y` are `J` for judgments, `O` for orders, `V` for opinions of the Court, `D` for decisions, `X` for rulings, `S` for seizure orders, `T` for third-party proceedings, `C` for AG opinions, and `P` for AG views. Some CELEX numbers also have a suffix with the format `(##)`, where the number is a counter (with preceding zeros) that differentiates between decisions of the same type related to the same case.

`decision_type` `string` The type of the decision. The possible values are (in decreasing order of frequency): `judgment`, `order`, `AG opinion`, `AG view`, `opinion`, `decision`, `seizure order`, `third-party proceedings`, and `ruling`. Note that not all types of decisions are possible at all courts. At the Court of Justice, all values are possible. At the General Court, the possible values are: `judgment`, `order`, `third-party proceedings` (rare), and `AG opinion` (rare). At the Civil Service Tribunal, the possible values are `judgment` and `order`.

`decision_date` `date` The date that the decision was published by the Court in the format `YYYY-MM-DD`.

`iuropa_procedure_id` `string` The IUROPA procedure ID. Uniquely identifies each legal procedure at the Court. IUROPA procedure IDs have the form `P:X`, where `X` is an abbreviation for the legal procedure. For procedures that have a corresponding proceeding suffix, the abbreviation matches the suffix. Otherwise, the abbreviation is an acronym for

the procedure name. The procedures `request for an opinion` and `ruling` use `OPINION` and `RULING`, respectively.

`procedure` `string` The legal procedure. The possible legal procedures are (in alphabetical order): `action for annulment`, `action for damages`, `action for failure to act`, `action for failure to fulfill obligations`, `appeal`, `appeal brought against a sanction`, `application for interim measures`, `application for interpretation of a judgment`, `application to intervene`, `application to set aside a judgment by default`, `arbitration clause`, `reference for a preliminary ruling`, `responsibility of a member of the institutions`, `review`, and `staff case`. See the `summary_procedures` table for more information.

`is_successful` `dummy` Whether the Court ruled that application with respect to the legal procedure was at least partially successful based on the merits.

`is_unfounded` `dummy` Whether the Court ruled that application with respect to the legal procedure was at least partially unfounded based on the merits.

`is_inadmissible` `dummy` Whether the Court ruled that application with respect to the legal procedure was at least partially inadmissible.

`is_unnecessary` `dummy` Whether the Court ruled that it was unnecessary to render a decision with respect to the legal procedure.

`is_interlocutory` `dummy` Whether the Court issued a temporary decision with respect to the legal procedure and reserved judgment on the application.

`no_disposition` `dummy` Whether the ruling of the Court does not have a disposition. There is no disposition when `procedure` is `reference for a preliminary ruling` or `arbitration clause`.

Assignments

All assignments of judges to judgments

Description

The `assignments` table includes data on which judges are assigned to the panel that is responsible for each judgment. There is one observation per judge per judgment. The `assignments` table also indicates which judge served as the Judge-Rapporteur, which is the judge who manages the proceeding and who writes the initial draft of the judgment.

We make a distinction between *chambers* and *panels*. Each Court has standing groups of judges called *chambers*. The composition of those chambers varies from session to session based on a complex rotation system. However, some chambers have more judges attached to them than can actually hear a case. For example, some chambers have six judges attached to them, but a maximum of five can hear any given case. We refer to the set judges from a *chamber* that actually hears a given case as a *panel*.

The data for the `assignments` table comes from several sources. For early judgments, we hand-code the data based on the text of the judgment. For later judgments, we scrape the data from the text of the judgment. We also collect data on the Judge-Rapporteur from EUR-Lex metadata. There are a small number of judgments where the text is not available, and the `assignments` table does not include observations for those judgments.

We cross-reference the data in several ways. First, we check that the Judge-Rapporteur identified in the EUR-Lex metadata matches the data we collect from the text of the decision and reconcile any discrepancies based on the text of the judgment. Second, we check that the Judge-Rapporteur is in the list of judges on the panel. Third, we check that each panel has an odd number of judges. Fourth, we check that the number of judges on

the panel is consistent with the formation of the Court, which we collect from the text of the judgment and from EUR-Lex metadata.

Note that the number of judges for most formations can vary from case to case. For example, some chambers can hear cases in 3-judge or 5-judge formations. The Full Court and the Grand Chamber also have quorum rules, which allows the number of judges to vary from case to case above a certain minimum. The size of the Court and the quorum rules have changed over time, so the number of judges in the `Full Court` and `Grand Chamber` formations can vary significantly over time.

The Judge-Rapporteur is available for all Court of Justice and Civil Service Tribunal judgments. However, at the General Court, judgments have not always indicated which judge was the Judge-Rapporteur. We code the Judge-Rapporteur whenever the text of the judgment or EUR-Lex metadata identifies the Judge-Rapporteur. At the General Court, data on the Judge-Rapporteur is available consistently from 2004. When information about the Judge-Rapporteur is not available, the `judge_rapporteur` dummy variable is coded `0` for all judges on the panel.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`court_id` `integer` An ID number that uniquely identifies each court. Coded `1` for the Court of Justice, `2` for the General Court, and `3` for the Civil Service Tribunal.

`court` `string` The court at which the case was lodged. The possible values are: `Court of Justice`, `General Court`, and `Civil Service Tribunal`.

`iuropa_case_id` `string` The IUROPA case ID number. Uniquely identifies each case. Coded based on `cjeu_case_id`. IUROPA case IDs have the format `CJEU:X:####:####`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, and the second number

is the number of the case (with preceding zeros). The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`cjeu_case_id` `string` The CJEU case ID number. Uniquely identifies each case. CJEU case IDs have the format `X-####/##`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, and the second number is the last two digits of the year. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it.

`case_year` `integer` The year that the case was lodged. Coded based on `cjeu_case_id`.

`case_number` `integer` The number of the case. Case numbers restart at `1` at the beginning of each year. Coded based on `cjeu_case_id`.

`iuropa_proceeding_id` `string` The IUROPA proceeding ID number. Uniquely identifies each proceeding. Coded based on `cjeu_proceeding_id`. IUROPA proceeding IDs have the format `CJEU:X:####:####[Y]`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, the second number is the number of the case (with preceding zeros), and `[Y]` is an optional suffix that distinguishes between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the suffix.

`cjeu_proceeding_id` `string` The CJEU proceeding ID number. Uniquely identifies each proceeding. CJEU proceedings IDs have the format `X-####/## Y`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, the second number is the last two digits of the year, and `Y` is an optional

suffix that differentiates between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it. Opinions of the Court and rulings use the prefixes `OPINION` and `RULING` instead of `C-`. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`.

`proceeding_suffix` `string` The proceeding suffix, if applicable. Coded based on `cjeu_proceeding_id`. This suffix indicates the legal procedure(s) associated with the proceeding. Coded `none` if `cjeu_proceeding_id` does not have a suffix. See the `summary_proceeding_suffixes` table for more information about the possible values and what each suffix means.

`proceeding_name` `string` The name of the proceeding. This is the name of the proceeding of record. If the Court has joined multiple proceedings together, the proceeding of record is the proceeding that was lodged first. The name of the case is presented exactly as it appears in the Registry. Coded `not available` for proceedings that do not have a record in the Registry.

`iuropa_decision_id` `string` The IUROPA decision ID number. Uniquely identifies each decision. IUROPA decision IDs have the format `CJEU:X:####:####:Y:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year of the case, the second number is the case number (with preceding zeros), `Y` is a letter that indicates the type of the decision, and the third number is the date of the document in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values of `Y` are based on CELEX descriptors: `J` for judgments, `O` for orders, `V` for opinions of the Court, `D` for decisions, `X` for rulings, `S` for seizure orders, `T` for third-party proceedings, `C` for AG opinions, and `P` for AG views. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the date.

cjeu_decision_id string The CJEU decision ID number. Uniquely identifies each decision. CJEU decision IDs are a combination of the ECLI number (**ecli**) and the date of the decision (**decision_date**). CJEU decision IDs have the format **ECLI:EU:X:####:####:#####**, where **X** is a letter that indicates the court that issued the decision, the first number is the year the decision was published, the second number is the document number assigned by the Court, and the third number is the date in the format **YYYYMMDD**. The possible values of **X** are **C** for the Court of Justice, **T** for the General Court, and **F** for the Civil Service Tribunal.

ecli string The ECLI number for the decision. Does not always uniquely identify each decision. ECLI numbers have the format **ECLI:EU:X:####:####**, where **X** is a letter that indicates the court that issued the decision, the first number is the year the decision was published, and the second number is a document number assigned by the Court. The possible values of **X** are **C** for the Court of Justice, **T** for the General Court, and **F** for the Civil Service Tribunal.

celex string The CELEX number for the decision. Coded **not applicable** for decisions that do not appear in EUR-Lex. Does not always uniquely identify each decision. CELEX numbers for CJEU decisions have the format **6####XY####**, where the first number is the year that the decision was published, **XY** is the CELEX descriptor for the decision, and the second number is the number of the case (with preceding zeros). The letter **X** indicates the court and the letter **Y** indicates the type of the decision. The possible values of **X** are **C** for the Court of Justice, **T** for the General Court, and **F** for the Civil Service Tribunal. The possible values for **Y** are **J** for judgments, **O** for orders, **V** for opinions of the Court, **D** for decisions, **X** for rulings, **S** for seizure orders, **T** for third-party proceedings, **C** for AG opinions, and **P** for AG views. Some CELEX numbers also have a suffix with the format **(##)**, where the number is a counter (with preceding zeros) that differentiates between decisions of the same type related to the same case.

decision_type string The type of the decision. The possible values are (in decreasing order of frequency): **judgment**, **order**, **AG opinion**, **AG view**, **opinion**, **decision**, **seizure order**, **third-party proceedings**, and **ruling**. Note that not all types of decisions are possible at all courts. At the Court of

Justice, all values are possible. At the General Court, the possible values are: `judgment`, `order`, `third-party proceedings` (rare), and `AG opinion` (rare). At the Civil Service Tribunal, the possible values are `judgment` and `order`.

`decision_date` `date` The date that the decision was published by the Court in the format `YYYY-MM-DD`.

`iuropa_judge_id` `string` The IUROPA judge ID for the judge. Coded based on `iuropa_judge_id` from the `judges` table.

`judge` `string` The name of the judge. Coded based on `last_name_label` from the `judges` table, which uniquely identifies judges with the same last name by including their first initial where necessary.

`is_judge_rapporteur` `dummy` Whether the judge served as the Judge-Rapporteur, who is the judge who manages the case and who writes the initial draft of the judgment.

Observers

All institutional observers in CJEU proceedings

Description

The `observers` table includes data on observers in CJEU proceedings. There is one observation per institutional observer per judgment concerning a reference for a preliminary ruling at the Court of Justice. The table covers every reference for a preliminary ruling in the `judgments` table. The `observer` table only covers institutional observers: EU member states, EU institutions, EU agencies, EFTA member states, and EFTA institutions.

The data for the `observers` table comes from EUR-Lex metadata. When EUR-Lex metadata is not available, we code the information by hand based on the text of the judgment.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`court_id` `integer` An ID number that uniquely identifies each court. Coded `1` for the Court of Justice, `2` for the General Court, and `3` for the Civil Service Tribunal.

`court` `string` The court at which the case was lodged. The possible values are: `Court of Justice`, `General Court`, and `Civil Service Tribunal`.

iuropa_case_id `string` The IUROPA case ID number. Uniquely identifies each case. Coded based on `cjeu_case_id`. IUROPA case IDs have the format `CJEU:X:####:####`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, and the second number is the number of the case (with preceding zeros). The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

cjeu_case_id `string` The CJEU case ID number. Uniquely identifies each case. CJEU case IDs have the format `X-####/##`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, and the second number is the last two digits of the year. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it.

case_year `integer` The year that the case was lodged. Coded based on `cjeu_case_id`.

case_number `integer` The number of the case. Case numbers restart at `1` at the beginning of each year. Coded based on `cjeu_case_id`.

iuropa_proceeding_id `string` The IUROPA proceeding ID number. Uniquely identifies each proceeding. Coded based on `cjeu_proceeding_id`. IUROPA proceeding IDs have the format `CJEU:X:####:####[Y]`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, the second number is the number of the case (with preceding zeros), and `[Y]` is an optional suffix that distinguishes between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the suffix.

`cjeu_proceeding_id` `string` The CJEU proceeding ID number. Uniquely identifies each proceeding. CJEU proceedings IDs have the format `X-####/## Y`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, the second number is the last two digits of the year, and `Y` is an optional suffix that differentiates between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it. Opinions of the Court and rulings use the prefixes `OPINION` and `RULING` instead of `C-`. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`.

`proceeding_suffix` `string` The proceeding suffix, if applicable. Coded based on `cjeu_proceeding_id`. This suffix indicates the legal procedure(s) associated with the proceeding. Coded `none` if `cjeu_proceeding_id` does not have a suffix. See the `summary_proceeding_suffixes` table for more information about the possible values and what each suffix means.

`proceeding_name` `string` The name of the proceeding. This is the name of the proceeding of record. If the Court has joined multiple proceedings together, the proceeding of record is the proceeding that was lodged first. The name of the case is presented exactly as it appears in the Registry. Coded `not available` for proceedings that do not have a record in the Registry.

`iuropa_decision_id` `string` The IUROPA decision ID number. Uniquely identifies each decision. IUROPA decision IDs have the format `CJEU:X:####:####:Y:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year of the case, the second number is the case number (with preceding zeros), `Y` is a letter that indicates the type of the decision, and the third number is the date of the document in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values of `Y` are based on CELEX descriptors: `J` for judgments, `O` for orders, `V` for opinions of the Court, `D` for decisions, `X` for rulings, `S` for seizure orders, `T` for third-party proceedings, `C` for AG opinions, and `P` for AG views. To distinguish

proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the date.

`cjeu_decision_id` `string` The CJEU decision ID number. Uniquely identifies each decision. CJEU decision IDs are a combination of the ECLI number (`ecli`) and the date of the decision (`decision_date`). CJEU decision IDs have the format `ECLI:EU:X:####:####:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year the decision was published, the second number is the document number assigned by the Court, and the third number is the date in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`ecli` `string` The ECLI number for the decision. Does not always uniquely identify each decision. ECLI numbers have the format `ECLI:EU:X:####:####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year the decision was published, and the second number is a document number assigned by the Court. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`celex` `string` The CELEX number for the decision. Coded `not applicable` for decisions that do not appear in EUR-Lex. Does not always uniquely identify each decision. CELEX numbers for CJEU decisions have the format `6####XY####`, where the first number is the year that the decision was published, `XY` is the CELEX descriptor for the decision, and the second number is the number of the case (with preceding zeros). The letter `X` indicates the court and the letter `Y` indicates the type of the decision. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values for `Y` are `J` for judgments, `O` for orders, `V` for opinions of the Court, `D` for decisions, `X` for rulings, `S` for seizure orders, `T` for third-party proceedings, `C` for AG opinions, and `P` for AG views. Some CELEX numbers also have a suffix with the format `(##)`, where the number is a counter (with preceding zeros) that differentiates between decisions of the same type related to the same case.

decision_type string The type of the decision. The possible values are (in decreasing order of frequency): judgment, order, AG opinion, AG view, opinion, decision, seizure order, third-party proceedings, and ruling. Note that not all types of decisions are possible at all courts. At the Court of Justice, all values are possible. At the General Court, the possible values are: judgment, order, third-party proceedings (rare), and AG opinion (rare). At the Civil Service Tribunal, the possible values are judgment and order.

decision_date date The date that the decision was published by the Court in the format YYYY-MM-DD.

iuropa_observer_id string The IUROPA actor ID for the observer. Coded based on iuropa_actor_id from the summary_actors table.

observer string The name of the observer. Coded based on actor from the summary_actors table.

observer_type string The type of the observer. Coded based on actor_type from the summary_actors table. The values that occur in the table are: EU member state, EU institution, EU body, EU decentralized agency, EFTA member state, and EFTA institution.

Citations

All citations in CJEU decisions

Description

The `citations` table includes data on citations in CJEU decisions. There is one observation per unique citation per decision. The `citations` table indicates each document that a decision cites, which paragraphs of the decision the citation occurs in, and which sub-parts of the cited document are referenced, if applicable. It also indicates the date, type, and author(s) of the cited document. The cited documents are identified by their CELEX numbers.

The data for the `citations` table comes from EUR-Lex metadata. EUR-Lex metadata on citations is not available for all judgments.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`court_id` `integer` An ID number that uniquely identifies each court. Coded `1` for the Court of Justice, `2` for the General Court, and `3` for the Civil Service Tribunal.

`court` `string` The court at which the case was lodged. The possible values are: `Court of Justice`, `General Court`, and `Civil Service Tribunal`.

iuropa_case_id `string` The IUROPA case ID number. Uniquely identifies each case. Coded based on `cjeu_case_id`. IUROPA case IDs have the format `CJEU:X:####:####`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, and the second number is the number of the case (with preceding zeros). The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

cjeu_case_id `string` The CJEU case ID number. Uniquely identifies each case. CJEU case IDs have the format `X-####/##`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, and the second number is the last two digits of the year. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it.

case_year `integer` The year that the case was lodged. Coded based on `cjeu_case_id`.

case_number `integer` The number of the case. Case numbers restart at `1` at the beginning of each year. Coded based on `cjeu_case_id`.

iuropa_proceeding_id `string` The IUROPA proceeding ID number. Uniquely identifies each proceeding. Coded based on `cjeu_proceeding_id`. IUROPA proceeding IDs have the format `CJEU:X:####:####[Y]`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the year that the decision was published, the second number is the number of the case (with preceding zeros), and `[Y]` is an optional suffix that distinguishes between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`. To distinguish proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the suffix.

`cjeu_proceeding_id` `string` The CJEU proceeding ID number. Uniquely identifies each proceeding. CJEU proceedings IDs have the format `X-####/## Y`, where `X` is a letter that indicates the court at which the case was lodged, the first number is the number of the case, the second number is the last two digits of the year, and `Y` is an optional suffix that differentiates between proceedings associated with the same case. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. For Court of Justice cases that were lodged before the creation of the General Court, the Court does not use a `C-` prefix, but the database always includes it. Opinions of the Court and rulings use the prefixes `OPINION` and `RULING` instead of `C-`. See the `summary_proceeding_suffixes` table for more information about the possible values of `Y`.

`proceeding_suffix` `string` The proceeding suffix, if applicable. Coded based on `cjeu_proceeding_id`. This suffix indicates the legal procedure(s) associated with the proceeding. Coded `none` if `cjeu_proceeding_id` does not have a suffix. See the `summary_proceeding_suffixes` table for more information about the possible values and what each suffix means.

`proceeding_name` `string` The name of the proceeding. This is the name of the proceeding of record. If the Court has joined multiple proceedings together, the proceeding of record is the proceeding that was lodged first. The name of the case is presented exactly as it appears in the Registry. Coded `not available` for proceedings that do not have a record in the Registry.

`iuropa_decision_id` `string` The IUROPA decision ID number. Uniquely identifies each decision. IUROPA decision IDs have the format `CJEU:X:####:####:Y:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year of the case, the second number is the case number (with preceding zeros), `Y` is a letter that indicates the type of the decision, and the third number is the date of the document in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values of `Y` are based on CELEX descriptors: `J` for judgments, `O` for orders, `V` for opinions of the Court, `D` for decisions, `X` for rulings, `S` for seizure orders, `T` for third-party proceedings, `C` for AG opinions, and `P` for AG views. To distinguish

proceedings related to opinions and rulings from other proceedings, we add `:V` for opinions and `:X` for rulings before the date.

`cjeu_decision_id` `string` The CJEU decision ID number. Uniquely identifies each decision. CJEU decision IDs are a combination of the ECLI number (`ecli`) and the date of the decision (`decision_date`). CJEU decision IDs have the format `ECLI:EU:X:####:####:#####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year the decision was published, the second number is the document number assigned by the Court, and the third number is the date in the format `YYYYMMDD`. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`ecli` `string` The ECLI number for the decision. Does not always uniquely identify each decision. ECLI numbers have the format `ECLI:EU:X:####:####`, where `X` is a letter that indicates the court that issued the decision, the first number is the year the decision was published, and the second number is a document number assigned by the Court. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal.

`celex` `string` The CELEX number for the decision. Coded `not applicable` for decisions that do not appear in EUR-Lex. Does not always uniquely identify each decision. CELEX numbers for CJEU decisions have the format `6####XY####`, where the first number is the year that the decision was published, `XY` is the CELEX descriptor for the decision, and the second number is the number of the case (with preceding zeros). The letter `X` indicates the court and the letter `Y` indicates the type of the decision. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The possible values for `Y` are `J` for judgments, `O` for orders, `V` for opinions of the Court, `D` for decisions, `X` for rulings, `S` for seizure orders, `T` for third-party proceedings, `C` for AG opinions, and `P` for AG views. Some CELEX numbers also have a suffix with the format `(##)`, where the number is a counter (with preceding zeros) that differentiates between decisions of the same type related to the same case.

decision_type (string) The type of the decision. The possible values are (in decreasing order of frequency): judgment, order, AG opinion, AG view, opinion, decision, seizure order, third-party proceedings, and ruling. Note that not all types of decisions are possible at all courts. At the Court of Justice, all values are possible. At the General Court, the possible values are: judgment, order, third-party proceedings (rare), and AG opinion (rare). At the Civil Service Tribunal, the possible values are judgment and order.

decision_date (date) The date that the decision was published by the Court in the format YYYY-MM-DD.

cited_celex (string) The CELEX number of the cited document. See celex from the decisions table for details.

cited_date (date) The date that the cited document was published in the format YYYY-MM-DD. Coded not available if EUR-Lex metadata does not indicate the date.

cited_type (string) The type of the cited document. Coded based on the CELEX sector for the document, which is indicated by the first number or letter of the CELEX number. The possible values are (in order of sector): treaty (sector 1), international agreement (sector 2), legislation (sector 3), complementary legislation (sector 4), working document (sector 5), case law (sector 6), parliamentary question (sector 9), other document in the OJ C series (sector C), and EFTA document (sector E).

cited_subtype (string) The subtype of the cited document. Coded based on the CELEX descriptor for the document, which is indicated by the two letters in the middle of the CELEX number. For example, the CELEX descriptor for a Court of Justice judgment is CJ and the corresponding value of cited_subtype is judgment (Court of Justice). There are 94 unique values.

`count_cited_authors` `integer` The number of authors of the cited document, according to EUR-Lex metadata. Coded based on `list_cited_authors`. Coded `0` when `list_cited_authors` is coded `not available`.

`list_cited_authors` `string` A list of the authors of the cited document, according to EUR-Lex metadata, separated by a semi-colon. Coded `not available` if EUR-Lex metadata does not indicate the authors.

`cited_detail` `string` The sub-part of the document referenced in the citation, if applicable. Coded `none` if the citation references the whole document, according to EUR-Lex metadata. Coded `other` if the EUR-Lex metadata could not be cleanly parsed.

`count_locations` `integer` The number of locations (i.e., paragraph numbers) in the decision where the citation occurs, according to EUR-Lex metadata. Coded based on `list_locations`.

`list_locations` `string` A list of the locations (i.e., paragraphs numbers) in the decision where the citation occurs, according to EUR-Lex metadata, separated by a comma. In decisions that have numbered paragraphs (such as some AG opinions), the numbers in `list_locations` correspond to those paragraph numbers. In decisions that do not have numbered paragraphs, `list_locations` uses the paragraphs numbers that are used in the EUR-Lex metadata.

Actors (summary)

All EU-related actors

Description

The `summary_actors` table includes data on all EU-related actors that appear in the IUROPA CJEU Database Platform. The table covers EU member states, EU bodies, EU executive agencies, EU decentralized agencies, EU interinstitutional services, EU corporate bodies, Euratom agencies, CFSP agencies, EFTA member states, and EFTA institutions.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`iuropa_actor_id` `string` The IUROPA actor ID. Uniquely identifies each EU-related actor. IUROPA actor IDs have the format `A:X`, where `X` is an abbreviation that indicates the actor.

`actor` `string` The name of each actor.

`actor_type` `string` The type of each actor. The possible types are: `EU member state`, `EU (general)`, `EU institution`, `EU body`, `EU executive agency`, `EU decentralized agency`, `EU interinstitutional service`, `EU corporate body`, `EU Euratom agency`, `EU CFSP agency`, `EU (other)`, `EFTA member state`, and `EFTA institution`.

start_year

integer

The start year for the actor. For member states, this is the year of accession. For institutions and agencies, this is the date of creation.

Formations (summary)

All formations of the Court

Description

The `summary_formation` table includes data on the formations of the Court. There is one observation per formation per court. The table covers the Court of Justice, the General Court, and the Civil Service Tribunal. The `summary_formation` table indicates the year of the first decision issued by each formation and the number of judges that the formation can have.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`court_id` `integer` An ID number that uniquely identifies each court. Coded `1` for the Court of Justice, `2` for the General Court, and `3` for the Civil Service Tribunal.

`court` `string` The court at which the case was lodged. The possible values are `Court of Justice`, `General Court`, and `Civil Service Tribunal`.

`iuropa_formation_id` `string` The IUROPA formation ID. Uniquely identifies each formation of the Court. IUROPA formation IDs have the format `F:X:Y`, where `X` indicates the Court and `Y` indicates the formation. The possible values of `X` are `C` for the Court of Justice, `T` for the General Court, and `F` for the Civil Service Tribunal. The

possible values for `Y` are `01` through `10` for numbered chambers, `FC` for Full Court, `GC` for Grand Chamber, `AC` for Appeals Chamber, and `SJ` for Single Judge.

`formation` `string` The name of the formation.

`list_panel_sizes` `string` The possible panel sizes for the formation, separated by commas.

`start_year` `integer` The year of the first decision issued by the formation.

Procedures (summary)

All legal procedures at the Court

Description

The `summary_procedures` table includes data on the legal procedures at the CJEU. There is one observation per procedure. The `summary_procedures` table includes the proceeding suffixes associated with the procedure, the courts that are relevant to the procedure, the pre-Lisbon and post-Lisbon case law directory codes that relate to each procedure, and the legal basis of each procedure.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`iuropa_procedure_id` `string` The IUROPA procedure ID. Uniquely identifies each legal procedure at the Court. IUROPA procedure IDs have the form `P:X`, where `X` is an abbreviation for the legal procedure. For procedures that have a corresponding proceeding suffix, the abbreviation matches the suffix. Otherwise, the abbreviation is an acronym for the procedure name. The procedures `request for an opinion` and `ruling` use `OPINION` and `RULING`, respectively.

`procedure` `string` The name of the procedure.

`proceeding_suffix` `string` The proceeding suffix that corresponds to the procedure. Coded `none` if the procedure does not correspond to a proceeding suffix.

`list_relevant_courts` `string` A list of the courts that handle each procedure, separated by a comma.

`pre_lisbon_directory_code` `string` The directory code in the pre-Lisbon case law directory that corresponds to the procedure. Coded `none` if there is not a corresponding directory code.

`post_lisbon_directory_code` `string` The directory code in the post-Lisbon case law directory that corresponds to the procedure. Coded `none` if there is not a corresponding directory code.

`legal_basis` `string` The legal basis for the procedure, referencing articles from the Treaty on the Functioning of the European Union (TFEU), the Statute on the Court of Justice of the European Union, the Rules of the Procedure of the Court of Justice, and the Rules of Procedure of the General Court. If the procedure has a legal basis in multiple articles or documents, they are separated by a semi-colon.

Document types (summary)

All document types for cited documents

Description

The `summary_document_types` table includes data on each type of document that can be cited in CJEU decisions. There is one observation per document type. The `summary_document_types` table indicates the possible type of each document, the possible subtype of each document, and the corresponding CELEX sector and descriptor for each type and subtype.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`celex_sector` `string` The CELEX sector for the document, which is the first number or letter of the CELEX number.

`celex_descriptor` `string` The CELEX descriptor for the document, which is the two letters in the middle of the CELEX number.

`document_type` `string` The type of the document. Coded based on the CELEX sector for the document. The possible values are (in order of sector): `treaty` (sector 1), `international agreement` (sector 2), `legislation` (sector 3), `complementary legislation` (sector 4), `working document` (sector 5), `case law`

(sector 6), `parliamentary question` (sector 9),
`other document in the OJ C series` (sector C), and `EFTA document` (sector E).

`document_subtype` `string` The subtype of the document. Coded based on the CELEX descriptor for the document. For example, the CELEX descriptor for a Court of Justice judgment is `CJ` and the corresponding value of `cited_subtype` is `judgment (Court of Justice)`. There are `94` unique values.

Proceeding suffixes (summary)

All proceeding suffixes

Description

The `summary_proceeding_suffixes` table includes data on each proceeding suffix used in CJEU proceeding IDs. There is one observation per suffix. The `summary_proceeding_suffixes` table indicates each possible suffix in CJEU proceeding IDs, the corresponding suffix for IUROPA proceeding IDs, and the legal procedure(s) indicated by each suffix.

Variables

`key_id` `integer` An ID number that uniquely identifies each observation.

`proceeding_suffix` `string` The unique proceeding suffixes used in `cjeu_proceeding_id`.

`proceeding_suffix` `string` The unique proceeding suffixes used in `iuropa_proceeding_id`.

`proceeding_type` `string` The type of proceeding according to the suffix.
