

TCSS CIC Board of Directors

A Guide to the Dataset

March 2026

1. Overview

186,338

Total Records

41,401

Unique Organisations

22

Variables

1992–2026

Financial Year Span

This dataset provides annual board-of-directors composition data for Community Interest Companies (CICs) registered with Companies House. It captures the size, turnover, diversity, and tenure characteristics of CIC boards across the full history of the CIC legal form.

The dataset is produced by the **UK Third and Civil Society Sector Database** project, which assembles open administrative data into research-ready panel datasets covering charities, community interest companies, and other civil society organisations.

Each row represents one organisation in one financial year, capturing board size, director turnover, gender diversity, nationality diversity, age composition, tenure, and founding board retention. The data is derived from Companies House officer records and aligned to financial years using the XBRL accounts filing history.

2. What are Board of Directors Records?

Community Interest Companies (CICs) are a legal form designed for social enterprises that use their assets and profits for the benefit of the community. CICs are subject to a statutory community interest test, an asset lock that prevents private extraction of value, and regulatory oversight by the CIC Regulator. Directors of CICs bear responsibility for maintaining the company's community purpose, making board composition a key indicator of organisational governance and capacity.

Like all UK limited companies, CICs are legally required to register their officers (directors and secretaries) with Companies House. This public register records each officer's appointment date, resignation date (if applicable), nationality, date of birth (month and year), and occupation. These records form the basis of this dataset.

Companies House Officers Data

The Companies House officers API provides structured data on every person who has served as a director or secretary of a registered company. Key fields include the date of appointment, date of resignation (where applicable), the officer's nationality, their date of birth (year and month only), and their stated occupation. This information is publicly available and updated as companies file changes.

From Officers to Board Panels

To produce this dataset, individual officer records are mapped to financial years using an inclusive overlap approach: a director is counted as serving in a given financial year if any part of their tenure overlaps with any part of that financial year. Officers are then aggregated to produce board-level summary metrics for each organisation in each year.

This transformation converts person-level appointment records into an organisation-year panel, where each row summarises the composition of the board during a specific financial year. Metrics include the number of directors, rates of appointment and resignation, gender and nationality proportions, age statistics, tenure measures, and founding board retention.

3. Dataset Contents

3.7

Mean Board Size

2.9 yrs

Median Tenure

48.8%

Mean Proportion Female

1992–2026

Financial Year Span

Field Descriptions

Field	Description	Type	Coverage
uid	Organisation identifier (format: GB-COH-{company_number})	Text	100.0%
coyno	Companies House company number (zero-padded to 8 digits)	Text	100.0%
fye	Financial year end date (YYYY-MM-DD)	Date	100.0%
fy	Financial year as numeric value	Numeric	100.0%
board_size	Number of directors serving during the financial year	Integer	100.0%
appointments_in_fy	Number of new director appointments in the financial year	Integer	100.0%
resignations_in_fy	Number of director resignations in the financial year	Integer	100.0%
turnover_rate	Board turnover rate (resignations ÷ board size)	Numeric	100.0%
prop_female	Proportion of directors classified as female (excludes unknown gender from denominator)	Numeric	99.5%
prop_non_british_irish	Proportion of directors with non-British/Irish nationality (excludes unknown)	Numeric	100.0%
age_mean	Mean age of directors at financial year end (approximate, based on birth year only)	Numeric	100.0%
age_median	Median age of directors at financial year end	Numeric	100.0%
tenure_mean_years	Mean director tenure in years at financial year end	Numeric	100.0%
tenure_median_years	Median director tenure in years at financial year end	Numeric	100.0%
founding_board_size	Number of directors on the founding board (appointed on the earliest appointment date)	Integer	100.0%
prop_founding_still_serving	Proportion of founding board members still serving in this financial year	Numeric	100.0%

Sample Data

The table below shows the first five rows of the dataset. Numeric values are rounded to two decimal places for display.

uid	coyno	fye	fy	board_size	appointments_in_fy	resignations_in_fy	turnover_rate	prop_female	prop_non_british_irish	age_mean	age_median	tenure_mean_year
GB-COH-00044259	00044259	1994-01-31	1994.00	9	1	3	0.33	0.10	0.10	65.44	65.00	2.31
GB-COH-00044259	00044259	1995-01-31	1995.00	11	5	2	0.18	0.00	0.08	54.73	57.00	2.21
GB-COH-00044259	00044259	1996-01-31	1996.00	13	4	5	0.38	0.00	0.07	52.33	53.00	2.31
GB-COH-00044259	00044259	1997-01-31	1997.00	12	3	4	0.33	0.00	0.08	51.09	50.00	2.31
GB-COH-00044259	00044259	1998-01-31	1998.00	9	1	0	0.00	0.11	0.00	50.12	49.00	2.61

Field Coverage

The following fields have less than 100% coverage. Missing values typically reflect organisations that had no resignations, appointments, or officers with identifiable characteristics in a given financial year.

prop_female

99.5%

4. Coverage & Completeness

The dataset covers 41,401 unique CICs across financial years 1992–2026, with 186,338 organisation–year records in total. The coverage of each field is shown below.

uid	coyno	fye	fy
100.0%	100.0%	100.0%	100.0%
board_size	appointments_in_fy	resignations_in_fy	turnover_rate
100.0%	100.0%	100.0%	100.0%
prop_female	prop_non_british_irish	age_mean	age_median
99.5%	100.0%	100.0%	100.0%
tenure_mean_years	tenure_median_years	founding_board_size	prop_founding_still_serving
100.0%	100.0%	100.0%	100.0%
prop_prior_experience	prop_prior_cic_experience	prop_prior_charity_experience	prop_prior_cso_experience
100.0%	100.0%	100.0%	100.0%
prop_serve_currently		prop_serve_currently_cic	
100.0%		100.0%	

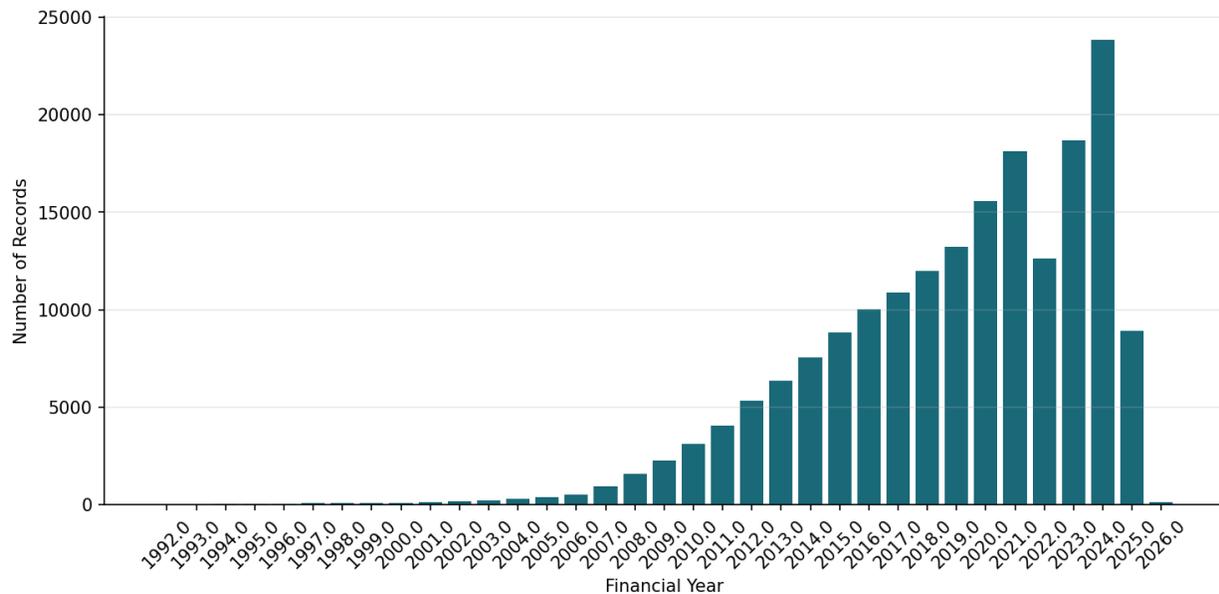


Figure 1. Number of organisation-year records by financial year. Source: UK Third Sector Database, Board of Directors dataset.

Unmatched Companies: An additional 29,824 CICs have officer data in Companies House but could not be matched to any financial year records in the XBRL accounts dataset. These are predominantly recently registered companies (76% registered in the 2020s) with a median lifespan of just 1.4 years — they either dissolved before filing accounts or are too new to have filed. All 29,824 are confirmed in the TCSS Organisation Register as CICs. See the Technical Annex for full details.

5. What Can You Learn?

This dataset enables a wide range of research into the governance and leadership of community interest companies. By providing annual board composition metrics across the full history of the CIC legal form, it supports both cross-sectional and longitudinal analyses of board dynamics.

Example research questions include:

- How has CIC board size changed over time?
- What is the gender composition of CIC boards, and has it improved?
- How long do founding directors remain on CIC boards?
- What is the relationship between board turnover and organisational age?
- How does nationality diversity on CIC boards compare across regions or sectors?

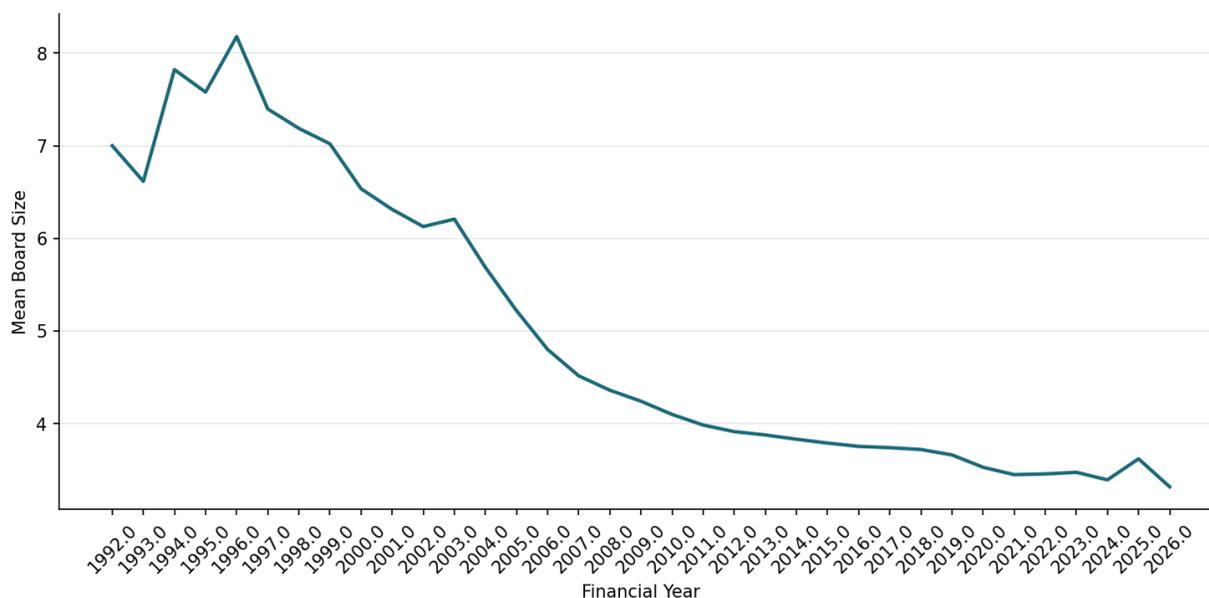


Figure 2. Mean board size by financial year. Source: UK Third Sector Database, Board of Directors dataset.

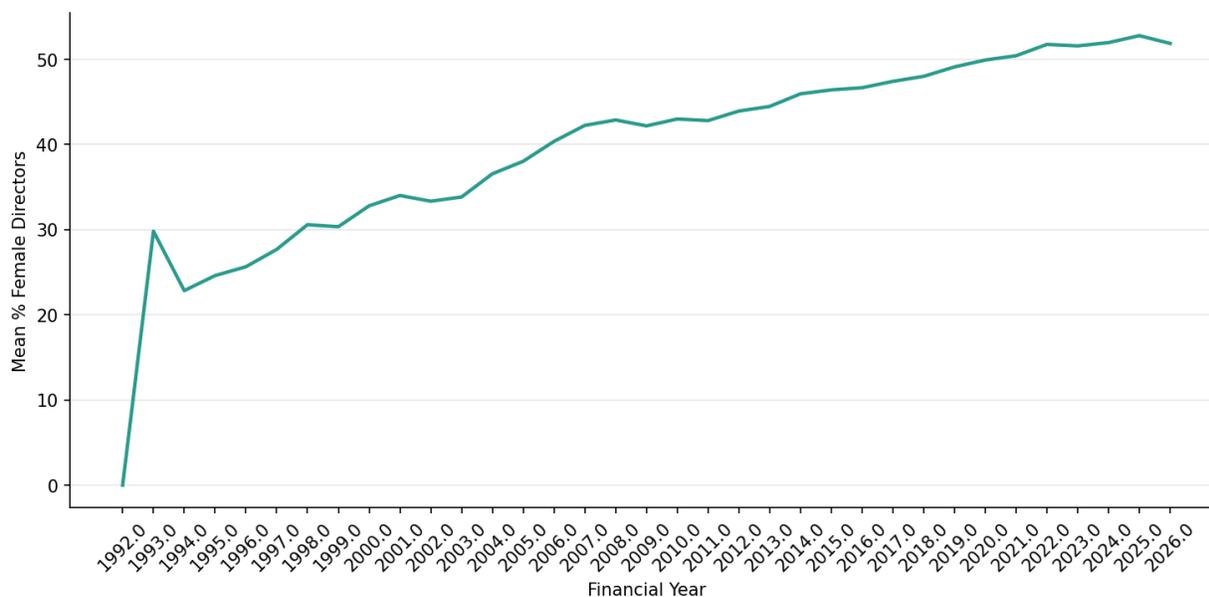


Figure 3. Mean proportion of female directors by financial year. Source: UK Third Sector Database, Board of Directors dataset.

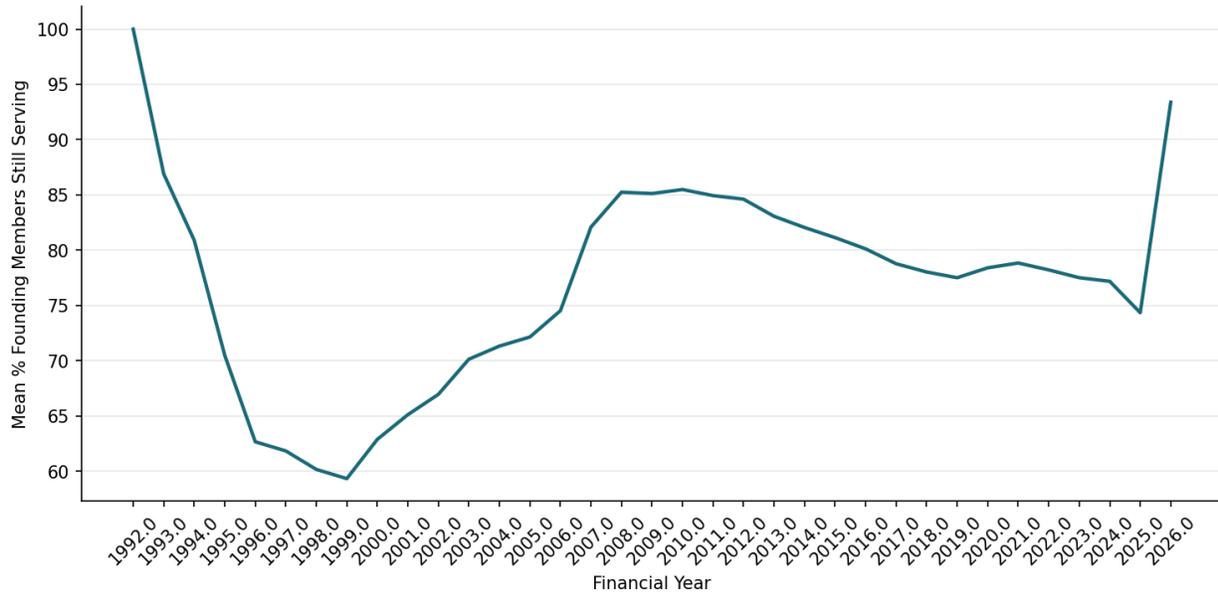


Figure 4. Mean proportion of founding board members still serving, by financial year. Source: UK Third Sector Database, Board of Directors dataset.

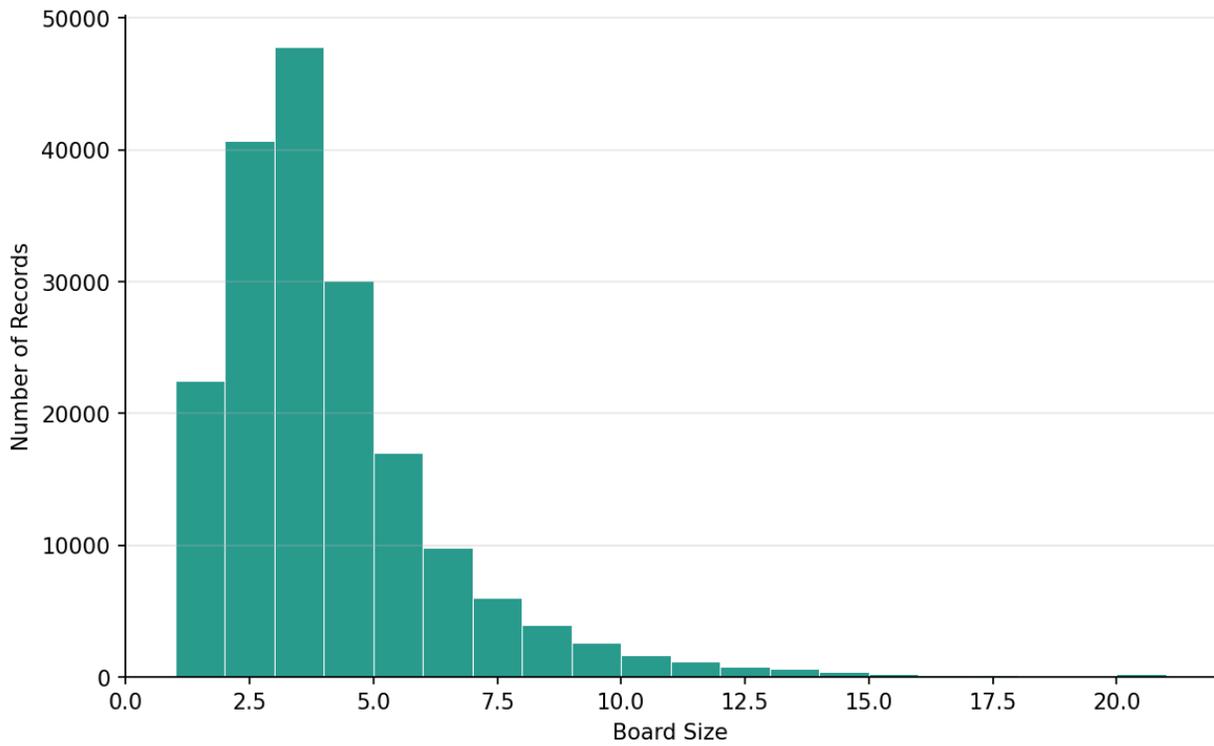


Figure 5. Distribution of board sizes across all organisation-year records (values above 20 grouped). Source: UK Third Sector Database, Board of Directors dataset.

Tip: Link this dataset to the TCSS Organisation Register using the `uid` field to add organisational characteristics such as registration date, company type, SIC codes, and registered address.

6. Limitations & Caveats

Gender Classification

Gender is inferred using a three-tier classification system: first by title (e.g. Mr, Mrs, Ms), then by forename matching against national name databases, and finally cross-validated using the UK Social Security Administration (SSA) name data. The proportion of directors with unknown gender is excluded from the denominator when computing `prop_female`. This approach may misclassify directors with non-Western names or gender-neutral forenames.

Age Approximation

Companies House records only the year of birth for company officers (month is also available but not used here). Age is therefore calculated as the financial year end year minus the birth year, which can be off by ± 1 year depending on the officer's actual birthday relative to the financial year end date.

Nationality

Nationality is self-reported by officers to Companies House and may not reflect current citizenship status. Many officer records have missing nationality data. For the purpose of computing `prop_non_british_irish`, British and Irish nationality variants are grouped as "domestic" and officers with unknown nationality are excluded from the denominator.

Financial Year Alignment

An inclusive overlap approach is used to assign officers to financial years: an officer is counted as serving in a financial year if any part of their tenure overlaps with any part of that financial year. This means an officer appointed on the last day of a financial year, or who resigned on the first day, is still counted as serving in that year.

Founding Board Definition

The founding board is defined as all directors appointed on the earliest appointment date for a given company. This date serves as a proxy for the incorporation date.

What is NOT in the Data

- **Individual officer names or identifiers** — data is aggregated to organisation level for disclosure protection.
- **Charities without Companies House registration** — The trustees of unincorporated charities can be downloaded from the CCEW data download portal, however this is a snapshot, not a longitudinal record: register-of-charities.charitycommission.gov.uk (<https://register-of-charities.charitycommission.gov.uk/en/register/full-register-download>).
- **Non-CIC companies** — this dataset covers Community Interest Companies only.
- **Companies without XBRL accounts filings** — 29,824 CICs with officer data but no matched financial year records are excluded. See the Coverage section for details.

7. Citation & Licence

Licence: This dataset is licensed under the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/) (CC BY 4.0). You are free to share, adapt, and build upon this data for any purpose, provided you give appropriate credit.

Suggested Citation

McDonnell et al. (2026). *TCSS CIC Board of Directors Dataset*. UK Third and Civil Society Sector Database. Available at: <https://uk-third-sector-database.github.io/data/>. Licensed under CC BY 4.0.

If you would like to learn more about this dataset and how it can be applied to your project or research programme, please contact research@brawdata.com.

8. Changelog

Version	Date	Description
1.0	March 2026	Initial release: 186,338 organisation-year records (1992–2026)

A1. Pipeline Architecture

The Board of Directors dataset is produced through a three-stage pipeline that transforms individual officer records from Companies House into an organisation-year panel of board composition metrics.

Officer Records → prepare_officers.py → align_to_financial_years.py → aggregate_board_summaries.py →

Board Panel

1. **Officer Data Preparation** — individual officer records are cleaned, standardised, and enriched with gender classifications and founding board flags
2. **Financial Year Alignment** — officers are mapped to the financial years during which they served, using an inclusive overlap rule against XBRL accounts data
3. **Board-Level Aggregation** — officer-year records are grouped by organisation and financial year to produce board-level summary metrics

A2. Officer Data Preparation

Date Standardisation

- Company numbers are zero-padded to 8 digits (preserving letter prefixes like SC, NI, OC)
- Appointment and resignation dates are parsed to standard format; `appointed_before` is used as a fallback when `appointed_on` is missing

Gender Classification

A three-tier priority system is used:

1. **Title-based (Tier 1)**: Mr → male; Mrs/Ms/Miss/Lady/Dame/Baroness → female; Dr/Prof/Rev/Sir/Lord → ambiguous (falls through)
2. **Forename-based (Tier 2)**: Uses the `gender_guesser` Python package; male/mostly_male → male; female/mostly_female → female; unknown/androgynous → unknown
3. **SSA Cross-Validation (Tier 2b)**: US Social Security Administration baby-name data provides an independent robustness check; stored in a separate column but does not override the primary classification

Priority: title wins over forename; the `gender_final` column reflects the title-then-name cascade.

Founding Board Identification

- The earliest `appointed_on` date per company serves as a proxy for incorporation date
- All officers appointed on that date are classified as founding board members
- `founding_board_size` is computed as the count of founding members per company

Forename Extraction

- Forenames are extracted from the Companies House appointments dataset by joining on `officer_id`
- Where multiple appointments exist for the same officer, the first non-null forename is retained
- Corporate officers (those flagged as `is_corporate_officer`) are identified and flagged

A3. Financial Year Alignment

Building the Organisation-Year Scaffold

- Financial year records are drawn from the XBRL accounts dataset (Companies House iXBRL filings)
- Deduplicated to unique company-number / financial-year-end pairs, retaining `uid`, `coyno`, `fy`, `fys` (FY start), and `fye` (FY end)

Overlap-Based Matching

Officers are matched to financial years using a vectorised inner join on company number, followed by an inclusive overlap filter:

- `appointed_on <= fye AND (resigned_on is missing OR resigned_on >= fys)`

This means an officer is counted as active in a financial year if any part of their tenure overlaps any part of the financial year period. For each officer-year match, the pipeline computes:

- `tenure_at_fye_years` : $(fye - appointed_on) \div 365.25$
- `appointed_in_fy` : True if `appointed_on` falls within `[fys, fye]`
- `resigned_in_fy` : True if `resigned_on` falls within `[fys, fye]`
- `age_at_fye` : `fye.year - dob_year` (approximate)

Unmatched Companies

29,824 CICs have officer data in Companies House but could not be matched to any financial year in the XBRL accounts dataset. Investigation confirms that all 29,824 are present in the TCSS Organisation Register and are flagged as CICs. They are predominantly recently registered companies: 76% were registered in the 2020s with a median lifespan of just 1.4 years. They have officer data because directors are registered immediately at incorporation, but no

XBRL accounts because they either dissolved before their first filing deadline or are too new to have filed. Only 31 (0.1%) have cross-register matches with other organisation types. These companies are logged for diagnostic purposes but are not included in the final dataset, as they have no financial year scaffold to align against.

A4. Board-Level Aggregation

The final stage groups officer-year records by organisation and financial year to produce one summary row per organisation per year.

Metric	Computation
board_size	Count of distinct officers per group
appointments_in_fy	Sum of appointed_in_fy flags
resignations_in_fy	Sum of resigned_in_fy flags
turnover_rate	resignations_in_fy ÷ board_size
prop_female	Female ÷ (male + female); unknowns excluded from denominator
prop_non_british_irish	Non-British/Irish ÷ known nationality; unknowns excluded. British/Irish variants include "British", "Irish", "British, Irish", "Irish, British"
age_mean / age_median	Standard mean/median of age_at_fye
tenure_mean_years / tenure_median_years	Standard mean/median of tenure_at_fye_years
founding_board_size	Constant per company (max per group)
prop_founding_still_serving	Count of founding members still active ÷ board_size

Individual-level counts (e.g., number of female directors, number of non-British directors) are excluded from the final output to reduce disclosure risk. Only proportions and structural counts are published.

A5. Linking to Other TCSS Datasets

The `uid` field provides a direct link to all other datasets in the TCSS ecosystem. All records in this dataset use the `GB-COH-` prefix format.

Dataset	Link Field	Description
Organisation Register	<code>uid</code>	Master register of civil society organisations with registration details, company type, SIC codes, and status
Charity Financial Records	<code>uid</code>	Longitudinal income and expenditure data for registered charities (CCEW, OSCR, CCNI)
Nonprofit Financial Records	<code>uid</code>	Companies House XBRL accounts data for CICs and other nonprofit companies
CIC 36 Community Interest Statements	<code>uid</code>	Beneficiary descriptions, activities, and surplus use statements from CIC incorporation forms
Procurement Records	<code>uid</code>	Public procurement contracts and grants awarded to civil society organisations

Tip: To link CIC board data with charity records where a CIC is also a registered charity, use the TCSS matches file to find cross-register UIDs. A company with UID `GB-COH-12345678` may have a corresponding charity UID such as `GB-CHC-200009`.

A6. Reproducibility

- Pipeline code is available in the project repository at github.com/uk-third-sector-database/tso-database-builder (<https://github.com/uk-third-sector-database/tso-database-builder>), in the `boards-of-directors` subdirectory
- Requires Python 3.12+, dependencies managed with `uv`
- Key dependencies: `pandas`, `gender_guesser`
- All paths and parameters are defined in a central `config.py` file

Running the full pipeline requires access to the Companies House officers API data and the XBRL accounts dataset. The officers API data is downloaded separately using the Companies House data collection scripts. The final board-of-directors dataset is available for download from the [project website](https://uk-third-sector-database.github.io/data/) (<https://uk-third-sector-database.github.io/data/>).

Document generated: March 2026

UK Third and Civil Society Sector Database — uk-third-sector-database.github.io (<https://uk-third-sector-database.github.io/>)