ADR UK
Data-driven change

ADR UK flagship datasets

User note: information in the right-hand column of each slide can be edited for your own use.

UKRI Economic and Social Research Council

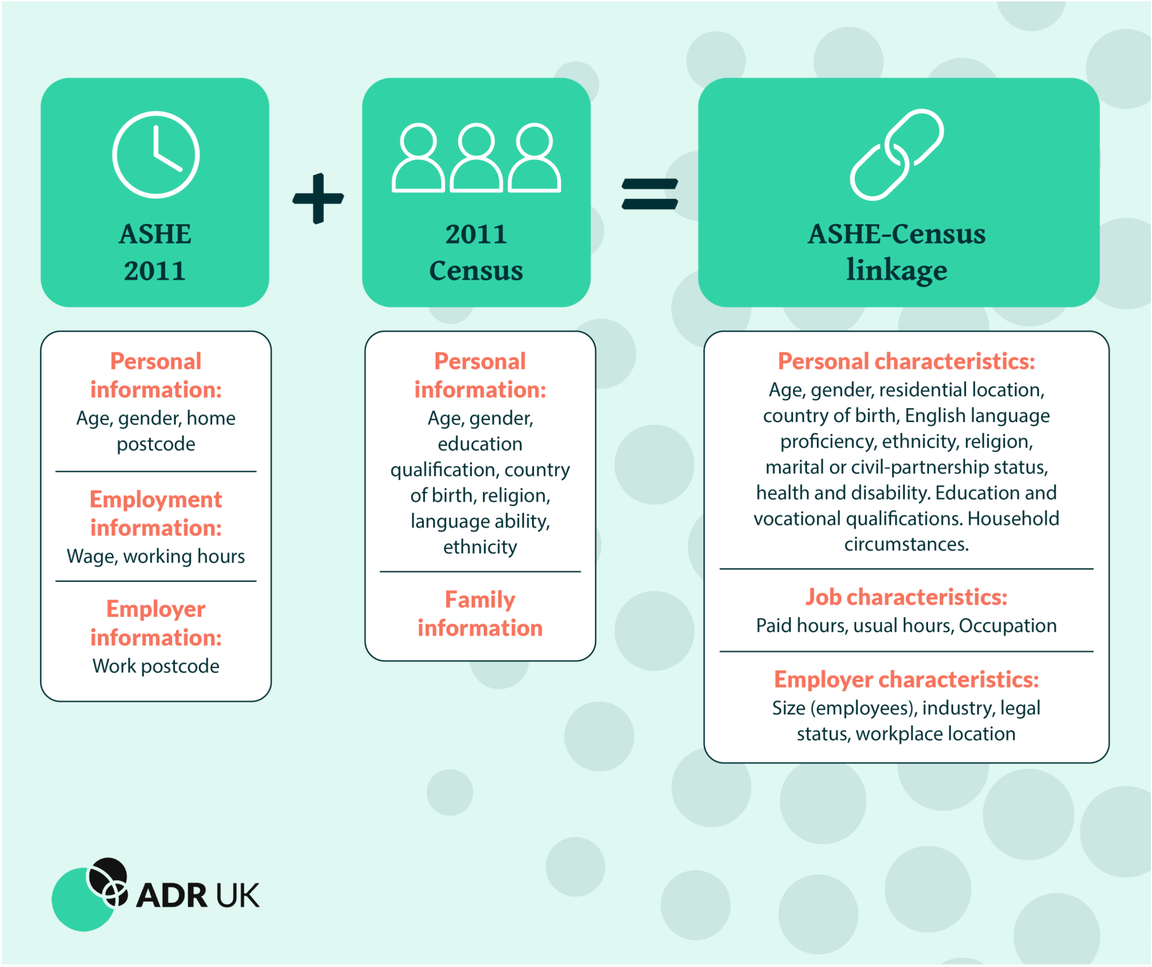
This deck will be regularly updated: last update October 2026 | Confidential material, not to be shared.

ADR UK flagship datasets:

1. Are selected for their potential to provide new data to **fill evidence gaps**
2. Have **significant research value** and wide appeal to researchers
3. Span a range of **policy-relevant themes**, including the economy, education, health, crime and justice
4. Have **approved legal gateways** for research
5. Are supported by a minimum standard of **documentation**

Find out more: Flagship datasets - ADR UK

Annual Survey of Hours & Earnings linked to 2011 Census (ASHE-2011 Census) – England and Wales



Dataset at a glance

- Information about levels, distribution and make-up of earnings and paid hours worked for employees in all industries and occupations
- 1997 – 2021
- ~175,000 records per year

Current research uses

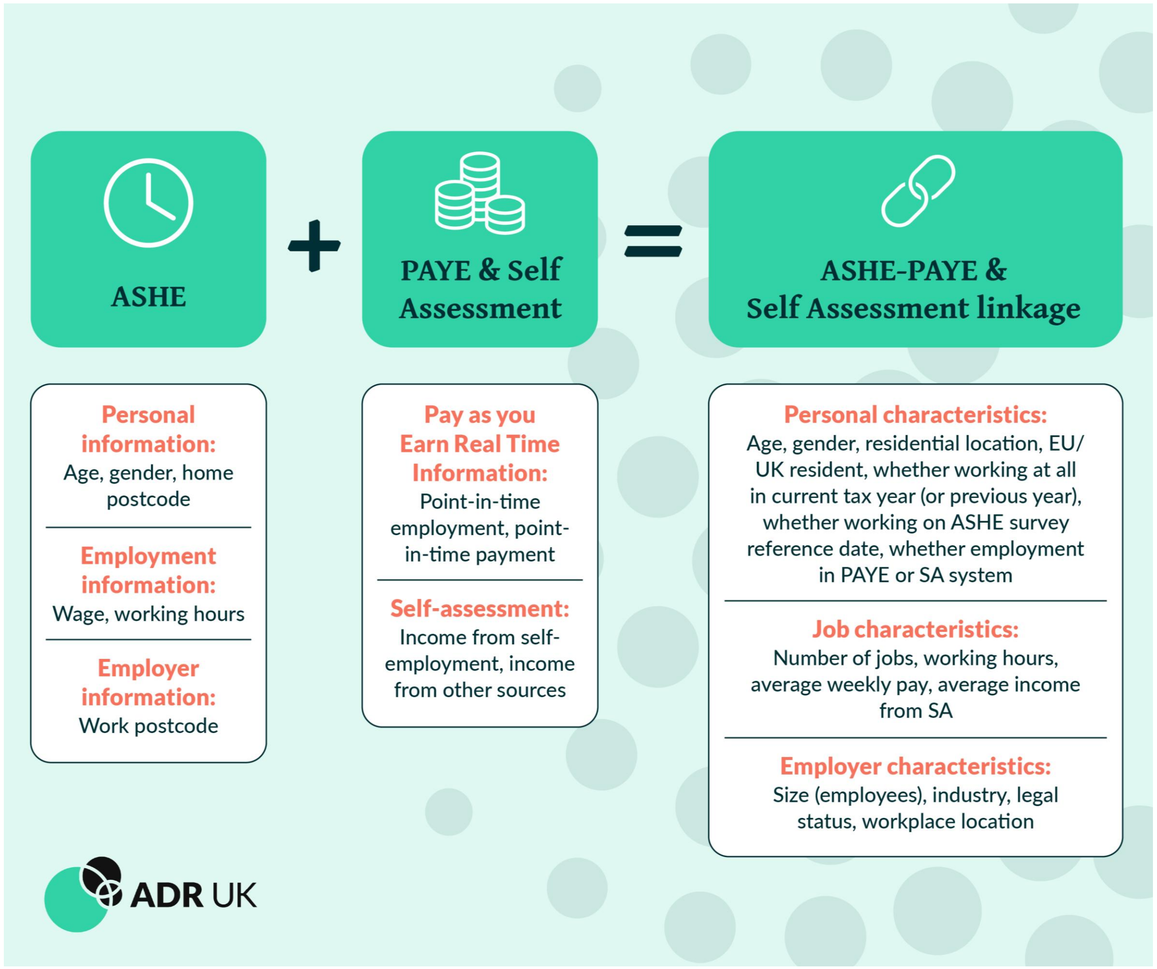
Projects using this dataset include:

- Measuring inequality-driven skills gaps in the UK labour market
- The nexus between immigration, integration and labour market outcomes
- The hidden cost of unpaid care: Evidence on the Care Pay Gap in England

Researcher support

- **Community:** ASHE Researcher Community allows researchers to connect, collaborate and share knowledge, insight and good practice
- **Training courses:** Wage and Employment Dynamics Team (WED) at UWE
- **Synthetic data:** v2011 (UKDS safeguarded)
- **Code:** ONS SRS (Stata)

Annual Survey of Hours and Earnings linked to PAYE and Self-Assessment data – England, Scotland and Wales



Dataset at a glance

- Comprehensive view of individuals' employment and income - self-employment and salaried work
- ASHE: 1997-2022/PAYE: 2015-19/SA: 2011-18
- ~130-180K individuals observed

Current research uses

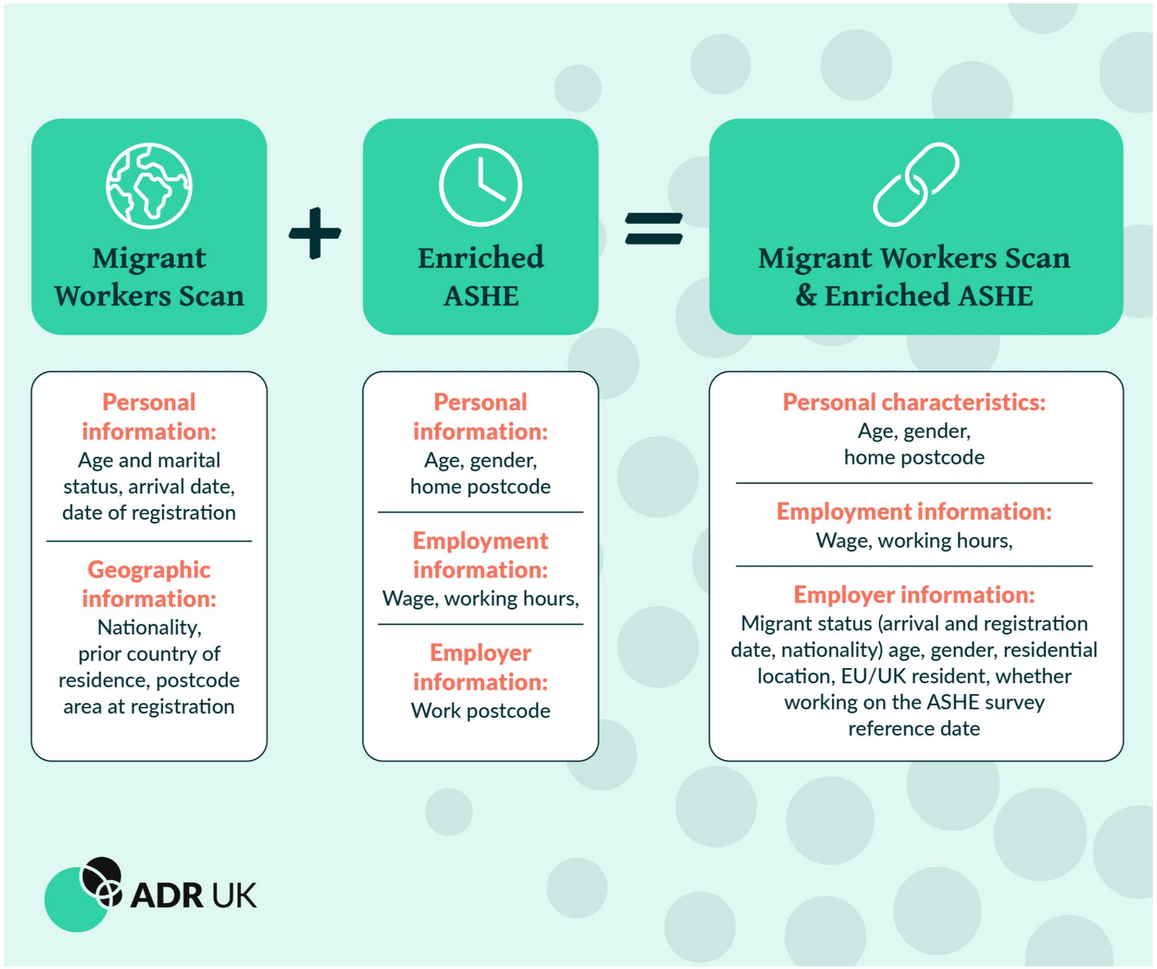
Projects using this dataset include:

- Uncovering patterns and policy implications of how employee jobs and self-employment are combined

Researcher support

- **Community:** ASHE Researcher Community allows researchers to connect, collaborate and share knowledge, insight and good practice
- **Training courses:** Wage and Employment Dynamics Team (WED) at UWE
- **Data Explained:** ASHE-HMRC Data Explained summarises experiences and learning from working with this dataset

Annual Survey of Hours and Earnings linked to Migrant Workers Scan (ASHE-MWS) – England, Scotland and Wales



Dataset at a glance

- Nationality and immigration details combined with pay, occupation and employment histories
- 1975 – 2022
- 156,200 persons in total

Current research uses

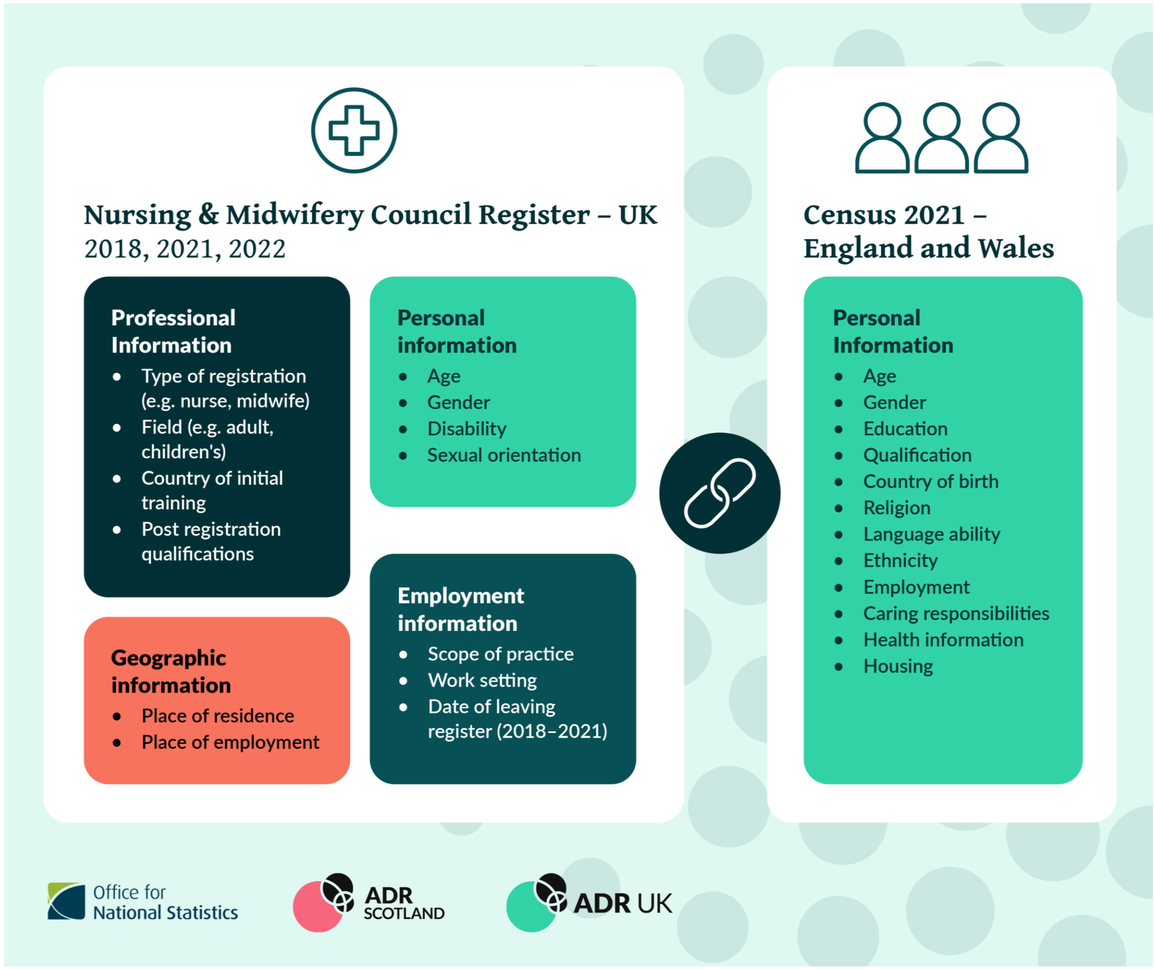
Projects using this dataset include:

- Why firms matter for immigrant pay gaps in the UK

Researcher support:

- **Community:** ASHE Researcher Community allows researchers to connect, collaborate and share knowledge, insight and good practice
- **Training courses:** Wage and Employment Dynamics Team (WED) at UWE
- **Data Explained:** ASHE-MWS Data Explained summarises experiences and learning from working with this dataset

Nursing and Midwifery Council Register linked to Census 2021 (NMC-Census) – England and Wales



Dataset at a glance

- Information provided by nurses, nursing associates, and midwives when they first join the register, and when they left - linked to Census
- NMC: 2018, 21, 22, Register leavers: 2018-21
- ~750,000 records in each year

Current research uses

Projects using this dataset include:

- The characteristics of early registrant leavers: who is most likely to leave the nursing and midwifery register early, and when
- Migration trends among nurses, nursing associates and midwives within the UK

Researcher support

- **Community:** Nursing and Midwifery Population Data Research Community - ADR Scotland platform for existing and prospective researchers, policymakers, practitioners, and voluntary and community sector organisations to learn, collaborate and innovate.

Data First: Cross-Justice System – England and Wales



Magistrates' courts defendant data
Timeframe: January 2011-March 2023

Crown Court defendant data
Timeframe: January 2013-March 2023

Family court data
Timeframe: January 2011-March 2023

Cross-justice system linking dataset
This is a lookup to connect records that refer to the same people across different MoJ datasets.

Civil court data
Timeframe: January 2012-January 2022

Prisoner custodial journey dataset
Timeframe: January 2001-March 2023

Offender assessment data
Timeframe: January 2011-December 2023

Probation dataset
Timeframe: January 2014-March 2023

Ministry of Justice

ADR UK

Dataset at a glance

- Connecting de-identified records referencing the same people across different justice datasets
- Unprecedented opportunity to understand cross-cutting questions about the overlaps and intersections between users of justice services

Current research uses

Projects using this dataset include:

- The impact of Home Detention Curfews on reoffending
- The criminal justice trajectories of perpetrators of violence against women and girls
- Understanding offender rehabilitation and supervision

Researcher support:

- **Community:** Crime and Justice Administrative Data Network a cross-institutional, cross-regional network open to anyone who uses, wants to use, or wants to learn from crime-related administrative data
- **Training courses:** Themed Workgroup: Courts & Justice– Datacise Open Learning
- **Synthetic data:** Data First SD available via UKDS

RAPID: Population interactions with employment, benefits and pensions – UK

Population	Demographics	Activities	Income	Disability and care
Covered	**Covered**	**Covered**	**Covered**	**Covered**
• UK coverage • Geography (Local Authority) • Children • Adults • Some household links • 2010/11 onwards • Resident / Not resident measure	• Birth • Age • Sex • Year of death • Alive post 05/04/2008 • Nationality (UK/EU/Rest of World) at National Insurance Number registration	• Benefits by type • Employment summary • Self-employment • Occupational and personal pensions • Benefit relationships and households	• Benefit income • Employment income • Self-employment profit • Pensions in payment (state pension and private)	• Disability condition • Component combinations • Care relationship • Mobility / daily living elements • Universal Credit health journey
Not covered	**Not covered**	**Not covered**	**Not covered**	**Not covered**
• All household links • Special populations (prisoners, foreign forces etc.)	• Ethnicity • Current nationality • Marital / relationship status • Gender	• Individual employment detail • Hours worked • Education activity and qualifications	• Investment income • Savings • Maintenance payments • Winter fuel • Outgoing, debt and housing costs	• Medical condition for Scottish benefits • Disabled but not on benefit



ADR UK

Dataset at a glance

- Annual summary of employment, benefit and pension activities, and the associated income
- 2010-11 to 2023-24
- 79-87 million person-records per tax year

Current research uses

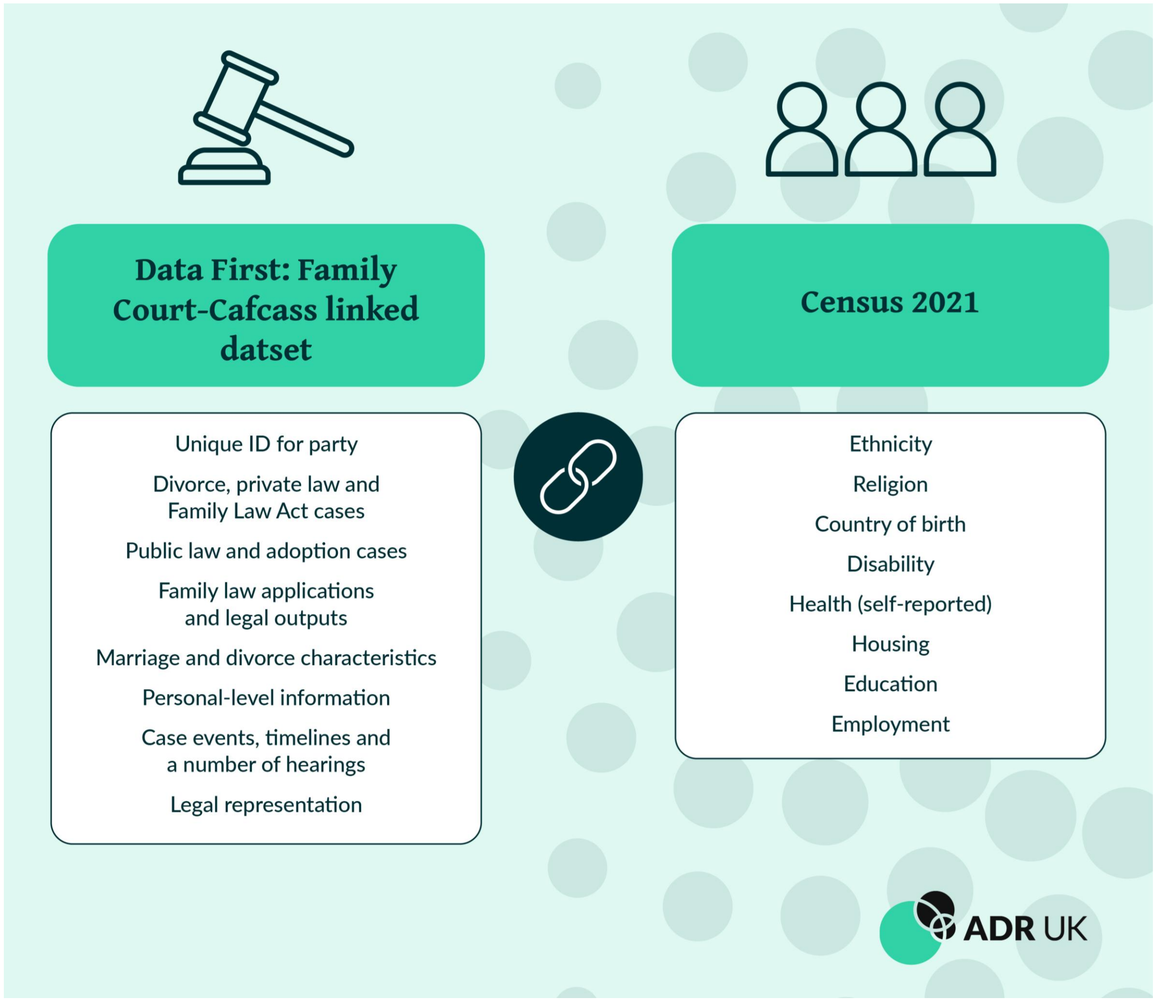
Projects using this dataset include:

- Labour market dynamics of older working-age people and the role of caring, ill-health, disability and local conditions
- The economic determinants of health-related benefit claims and health-related inactivity
- Economic inactivity in the UK

ADR UK Academic Lead team for Better Economic Data

Working with the Department for Work and Pensions (DWP) and His Majesty's Revenue and Customs (HMRC) the new Academic Lead team (Better Economic Data) has now been commissioned to support this and other economic datasets.

Data First: Family Court linked to Cafcass and Census 2021 – England and Wales



Dataset at a glance

- Links Ministry of Justice family court records with Cafcass data and Census 2021 demographic and household characteristics
- Timeframe: Family court: 2011–2021; Cafcass: 2007–2023; Census: 2021
- Records: Covers individuals and families who have interacted with the family justice system in England and Wales

Current research uses

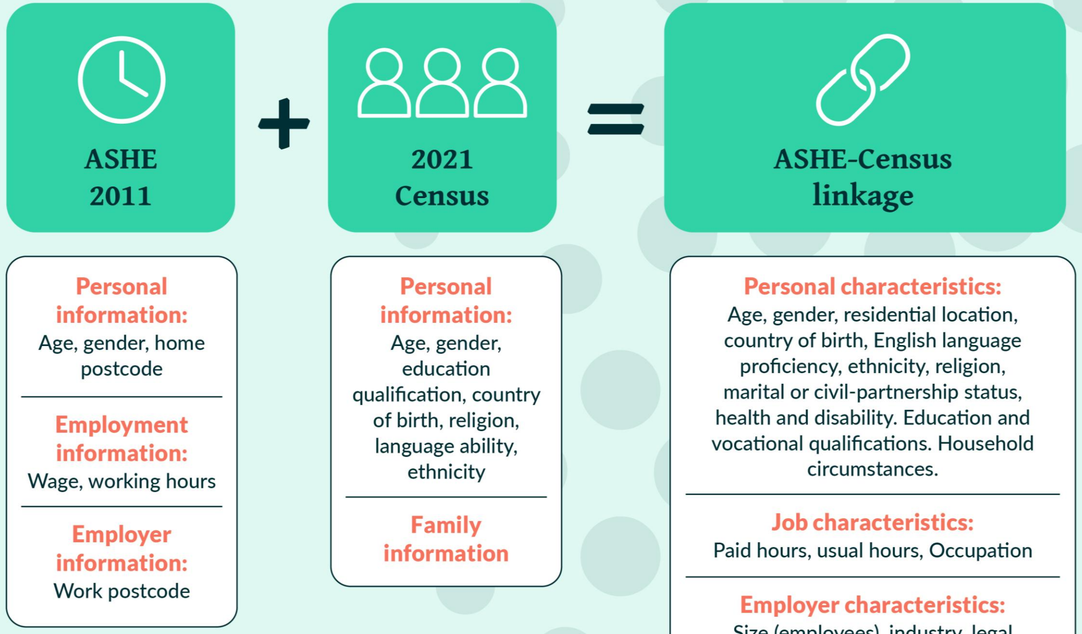
Projects using this dataset include:

- Domestic abuse proceedings, case overlaps and pathways in family courts
  Life course patterns and court experiences of young mothers in recurrent care proceedings

Dataset resources

- **User guide:** Ministry of Justice: Data First
- **Data dictionary:** Cafcass England (CAFE) – SAIL
- **Data dictionary:** Census 2021 dictionary
- **Dataset webinar:** Research Fellowships: Data First Family Court linked to Cafcass and Census 2021 dataset briefing

Annual Survey of Hours and Earnings (ASHE) linked to Census 2021 – England and Wales (coming soon)



Dataset at a glance

- Links de-identified annual individual earnings, hours worked and employer details with 2021 Census demographics, education and household information
- Timeframe: 1999–2022 (ASHE) and 2021 (Census 2021)
- Records: ~150,000 jobs per year in ASHE

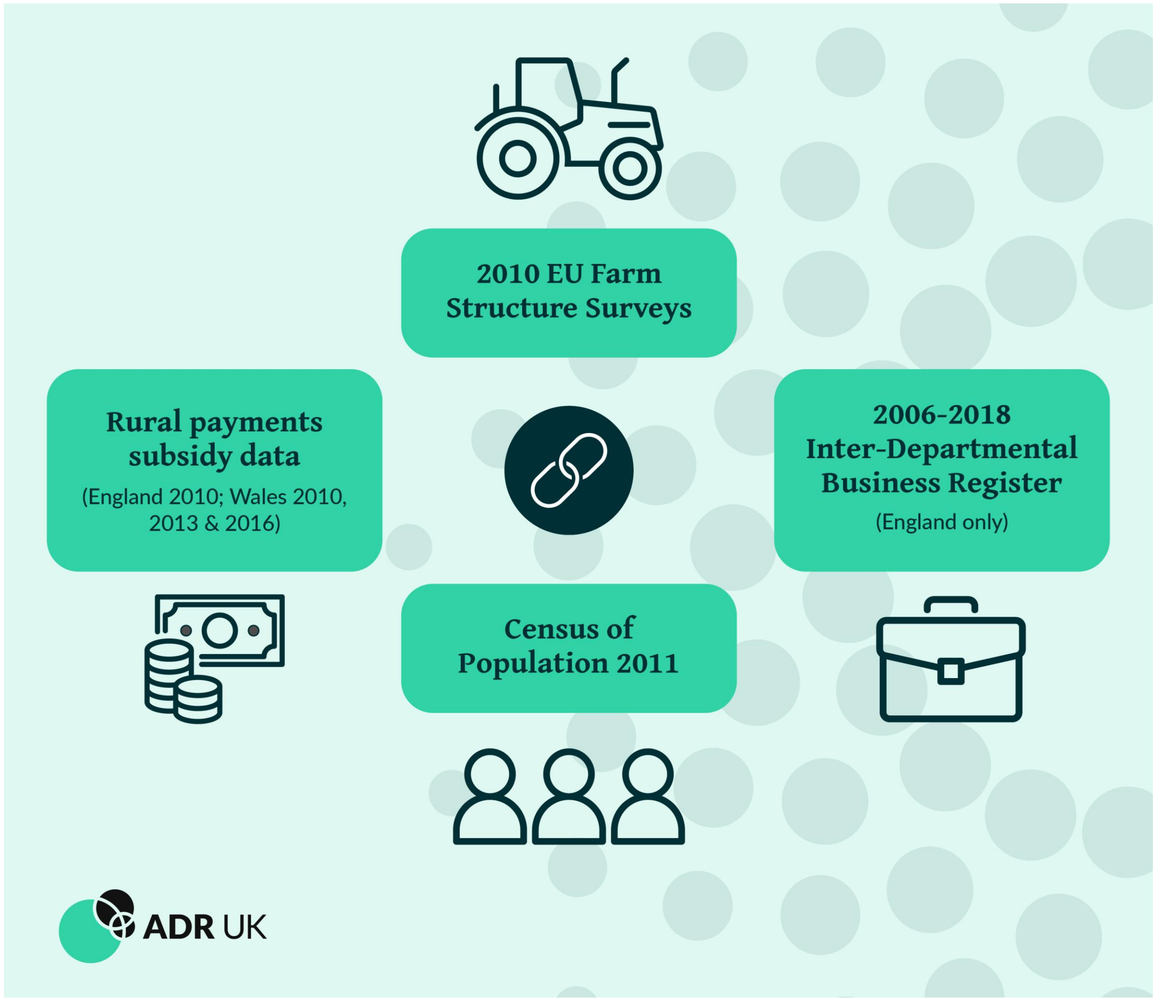
Research themes

- Career earnings trajectories across gender, disability and ethnicity
- Progression patterns and characteristics associated with low-paid employment
- Impact of household composition and migration status on labour market outcomes

Dataset resources

- **Beta versions of the user guide, data dictionary and synthetic dataset** are available on the WED website.
- **Contact:** wagedynamics@uwe.ac.uk

Administrative Data | Agricultural Research Collection (AD|ARC) 2011: England and Wales



Dataset at a glance

- Links farm business activities, land use and subsidy payments with de-identified 2011 Census data for farming households and rural control groups
- Timeframe: 2006–2018 (varies by dataset)
- Records: England 2011: ~170k individuals (~60k farms); Wales 2011: ~50k individuals (~18k farms)

Current research uses

Projects using this dataset include:

- Comparing farm and non-farm rural household structures
- Farm household resilience, income sources and non-farm earnings
- Socio-demographic profiling and health outcomes of farming families

Dataset resources

- **User guides and metadata:** AD|ARC Project Documentation
- **Project website:** AD|ARC Official Website

Administrative Data | Agricultural Research Collection (AD|ARC) 2021 – England and Wales



Dataset at a glance

- Farming business data linked with de-identified 2021 Census data for farming households and non-farming rural control groups
- Timeframe: 2010–2022 (varies by dataset)
- Records: England 2021: ~240k individuals (~90k farms); Wales 2021: ~40k individuals (~14k farms)

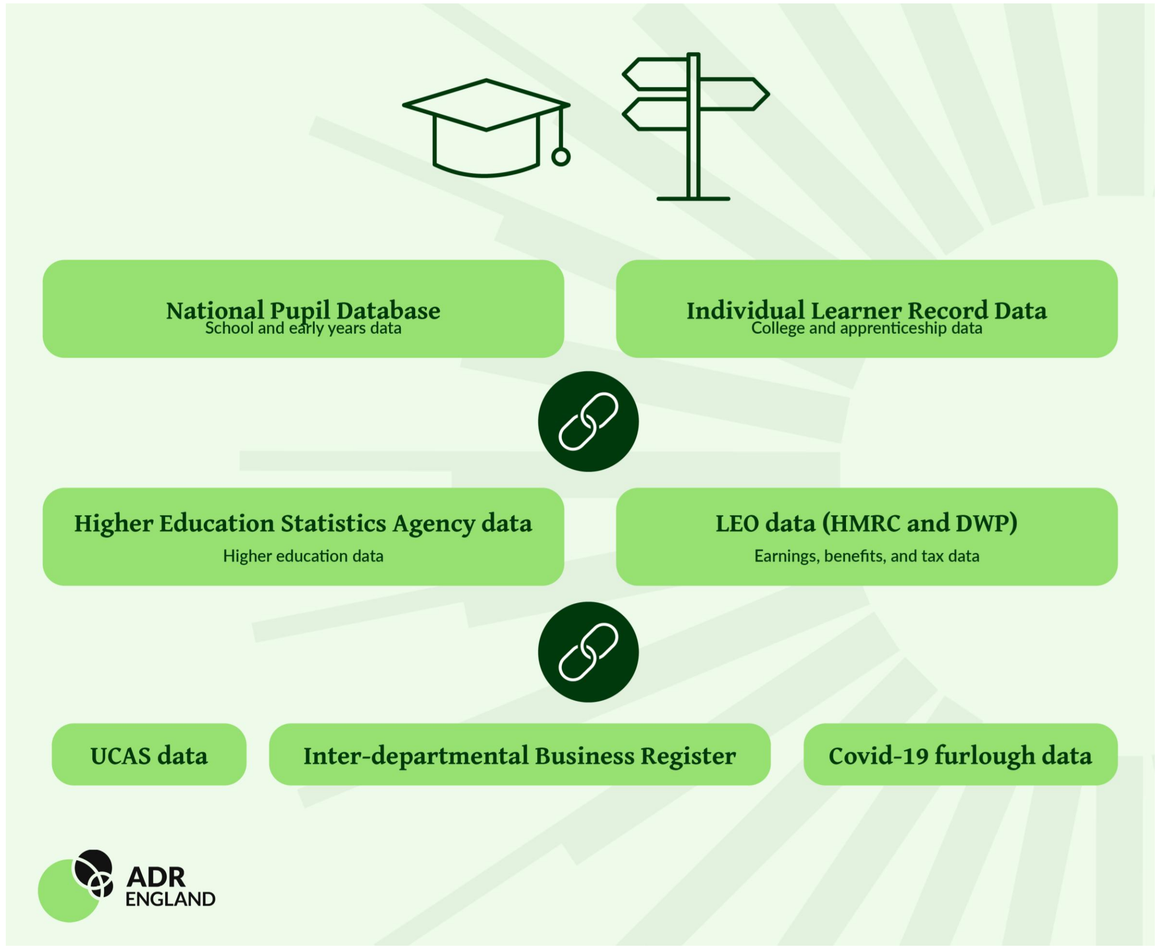
Research themes

- Comparing farm and non-farm rural household structures
- Farm household resilience, income sources and non-farm earnings
- Socio-demographic profiling and health outcomes of farming families

Dataset resources

- **User guides, metadata and synthetic data:** AD|ARC User Guides
- **Website:** AD|ARC Official Website

Longitudinal Education Outcomes (LEO) – England



Dataset at a glance

- De-identified information: characteristics, education, employment, benefits, earnings
- January 1995 – July 2021
- ~38 million records

Current research uses

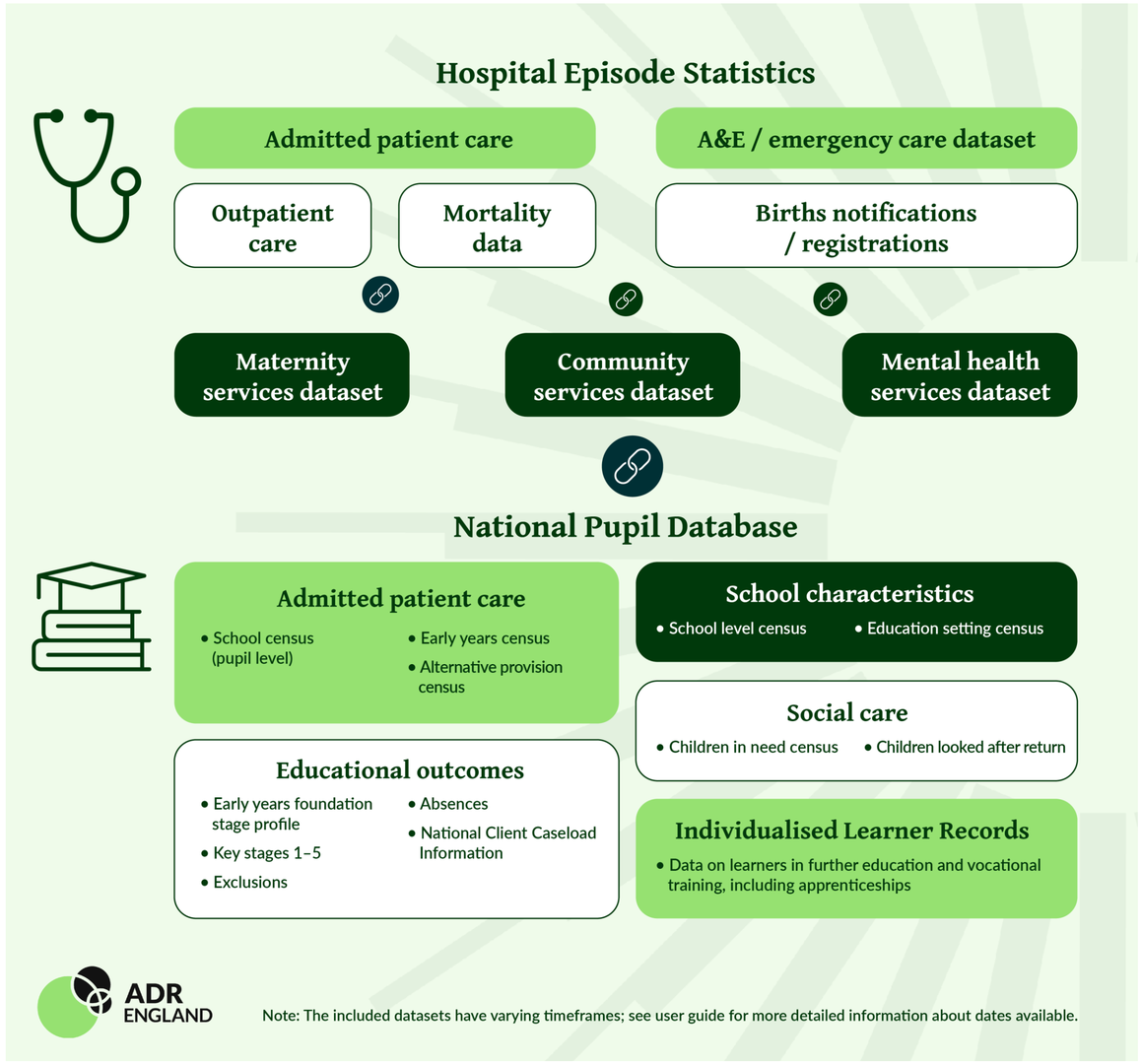
Projects using this dataset include:

- The impact of unaffordable housing on intergenerational mobility
- Educational pathways for young people with special educational needs and disabilities in England
- Investigating STEM readiness, inclusion, and economic returns

Researcher support

- **Community:** LEO Researcher Community supports data preparation and analysis and hosts a library of resources and information
- **Training courses:** UCL Centre for Education Policy and Equalising Opportunities (CEPEO)
- **Synthetic data:** V1, 95-2021 (UKDS safeguarded)
- **Code:** ONS SRS (SQL)

Education and Child Health Insights from Linked Data (ECHILD) – England



Dataset at a glance

- Data on pupil and school characteristics, educational outcomes, and social care - linked to healthcare data
- 1 September 1984 – 31 August 2022
- ~21 million records

Current research uses

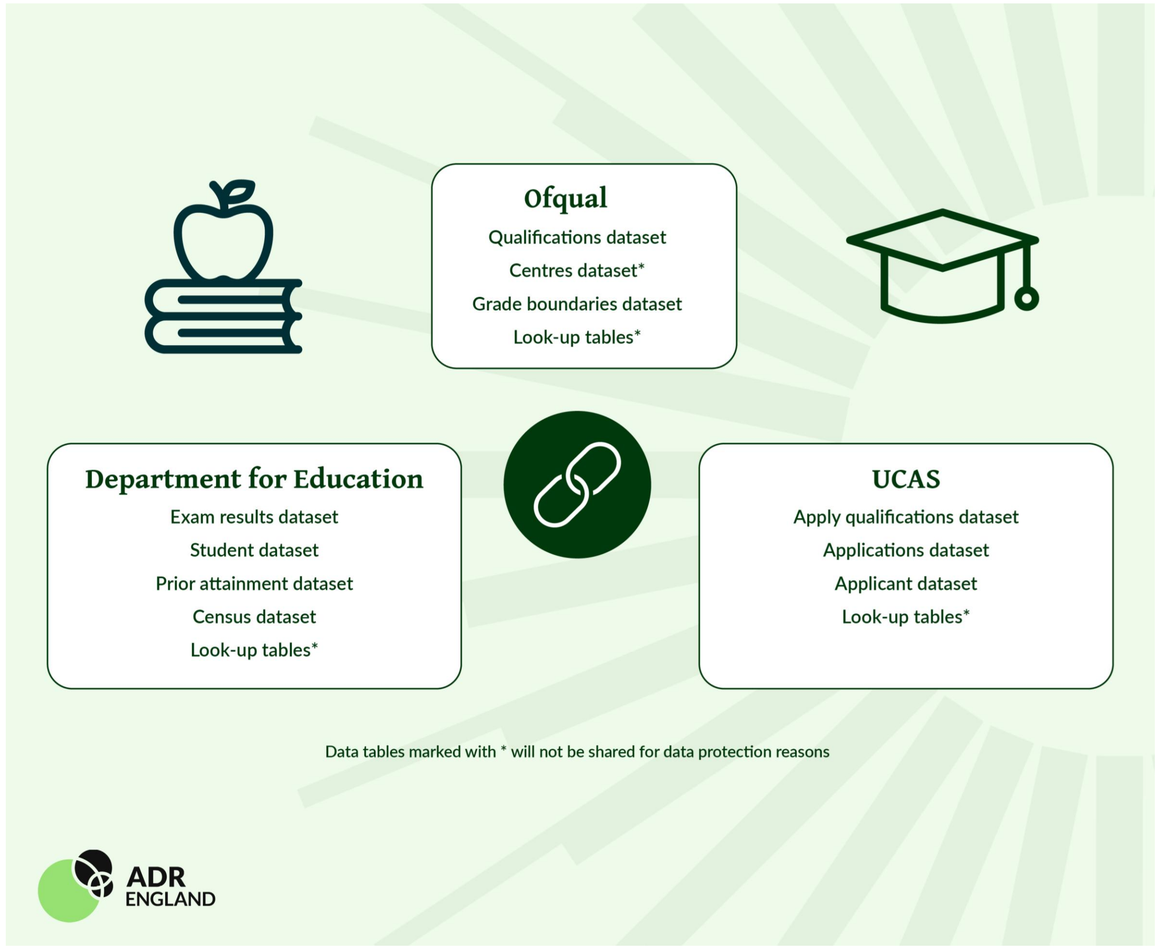
Projects using this dataset include:

- Health-related outcomes, alternative provision, and exclusion among pupils with neurodivergent special educational needs
- Women's mental illness in pregnancy: Exploring contact with secondary mental health services and links with offspring health and education outcomes
- Siblings in contact with children's social care

Researcher support

- **Website:** ECHILD
- **Community:** UCL-ECHILD Discussions a place to ask questions about ECHILD data and to raise issues or to start discussions.
- **Training courses:** UCL-Instats: Intro to ECHILD
- **Synthetic data:** Training Synthetic Data
- **Code:** GitHub, Phenotype Code List Repository

Grading and Admissions Data for England (GRADE)



Dataset at a glance

- Links English GCSE and A-level exam grades with de-identified student socio-demographic characteristics, school history and university application records
- Timeframe: Academic year 2016/17 to 2021/22
- Records: ~4 million records

Current research uses

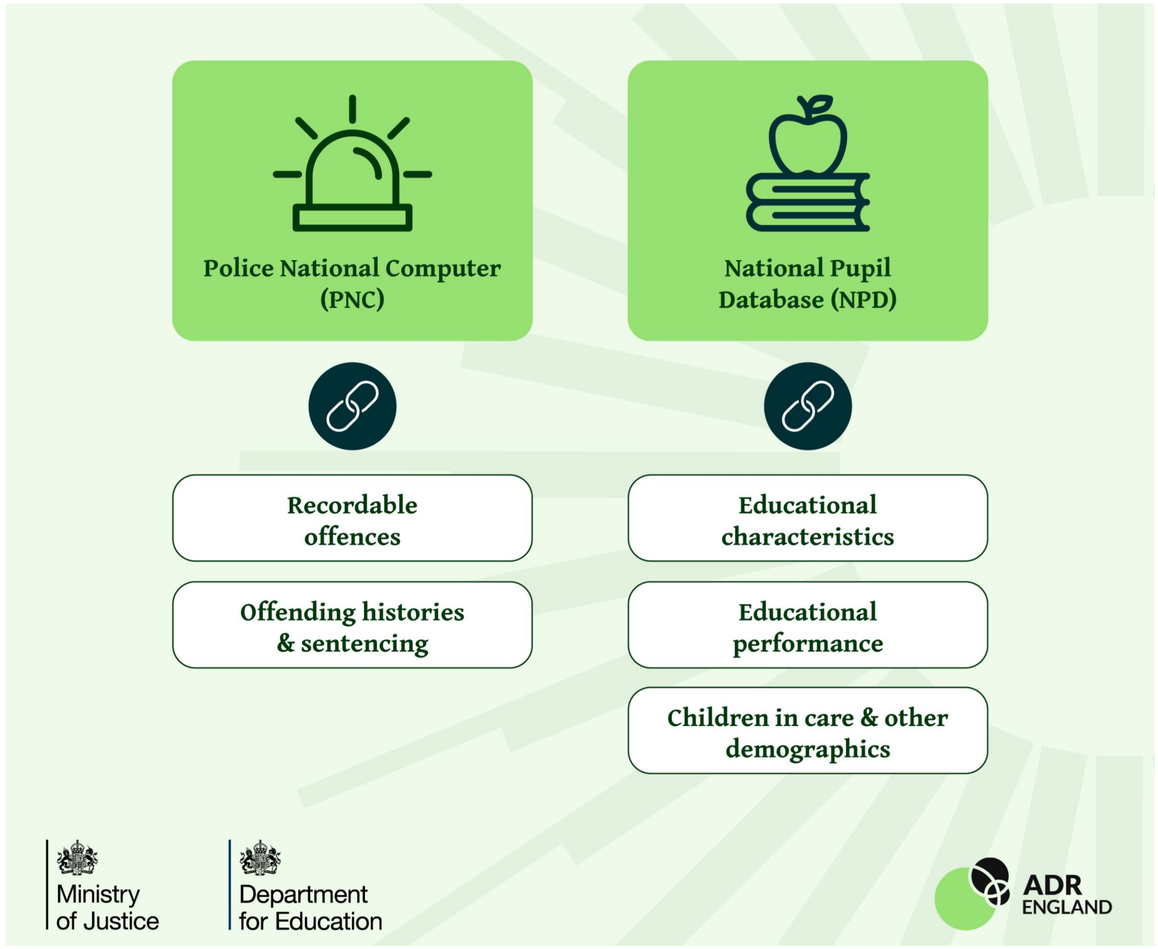
Projects using this dataset include:

- Quantifying pandemic-driven A-level assessment modifications and higher education admission outcomes
- Discrepancies in teacher grade predictions across student sociodemographic groups and university enrolment impacts

Dataset resources

- **User guide:** Grading and Admissions Data for England (GRADE) user guide
- **Data dictionary:** Data specifications for GRADE data sharing
- Synthetic data available on request; contact data.sharing@ofqual.gov.uk for more information

Ministry of Justice (MoJ) & Department for Education (DfE) linked dataset – England



Dataset at a glance

- De-identified criminal court records, prison custodial journeys and police cautioning data linked with school pupil characteristics, educational attainment and social care histories in England
- Timeframe: 2000–2025 (Police National Computer, MoJ) and pupils born after 31 August 1985 (National Pupil Database, DfE)
- Records: PNC: 1.65 million records; NPD: 15 million records

Current research uses

Projects using this dataset include:

- Education and social care predictors of offending trajectories
- Impact of child welfare service interventions on educational and criminal justice outcomes
- Intersections between care experiences and ethnicity in criminal justice involvement

Dataset resources

- **A user guide for this dataset is available through** Data First: An Introductory User Guide
- **Further information on all Data First linked datasets is** available from the MoJ
- **A public engagement case study about this dataset is also available through the** ADR UK website

Scottish Longitudinal Education Outcomes (LEO) – University extract



Higher Education Statistics (HESA) data
Higher Education delivered at Scottish Universities data

LEO data (HMRC and DWP)
Earnings, benefits and tax data

Inter-Departmental Business Register (IDBR)



Linked – but not yet shared for research

Scottish Funding Council (SFC) data
Higher and Further Education delivered at Scottish Colleges data

Skills Development Scotland (SDS) data
Modern apprenticeships data

Covid-19 furlough data

Dataset at a glance

- De-identified information on students who have attended Scottish Higher Education Institutions
- Links education, employment, earnings and benefits data

Research topics

- Earnings and employment outcomes for different institutions and subjects
- Differences in graduate outcomes by background characteristics, including care experience
- Industry destinations of graduates compared with Scotland's labour-market needs

Researcher support

- **Community:** A LEO community for the England dataset is co-ordinated by the Department for Education. Email ADRScotland@gov.scot to register interest in a Scotland-focused LEO community.

Looked After Children Longitudinal Dataset – Scotland



Dataset at a glance

- De-identified information on episodes of care, placements and the legal reason for children being looked after by local authorities in Scotland
- 72,000 children

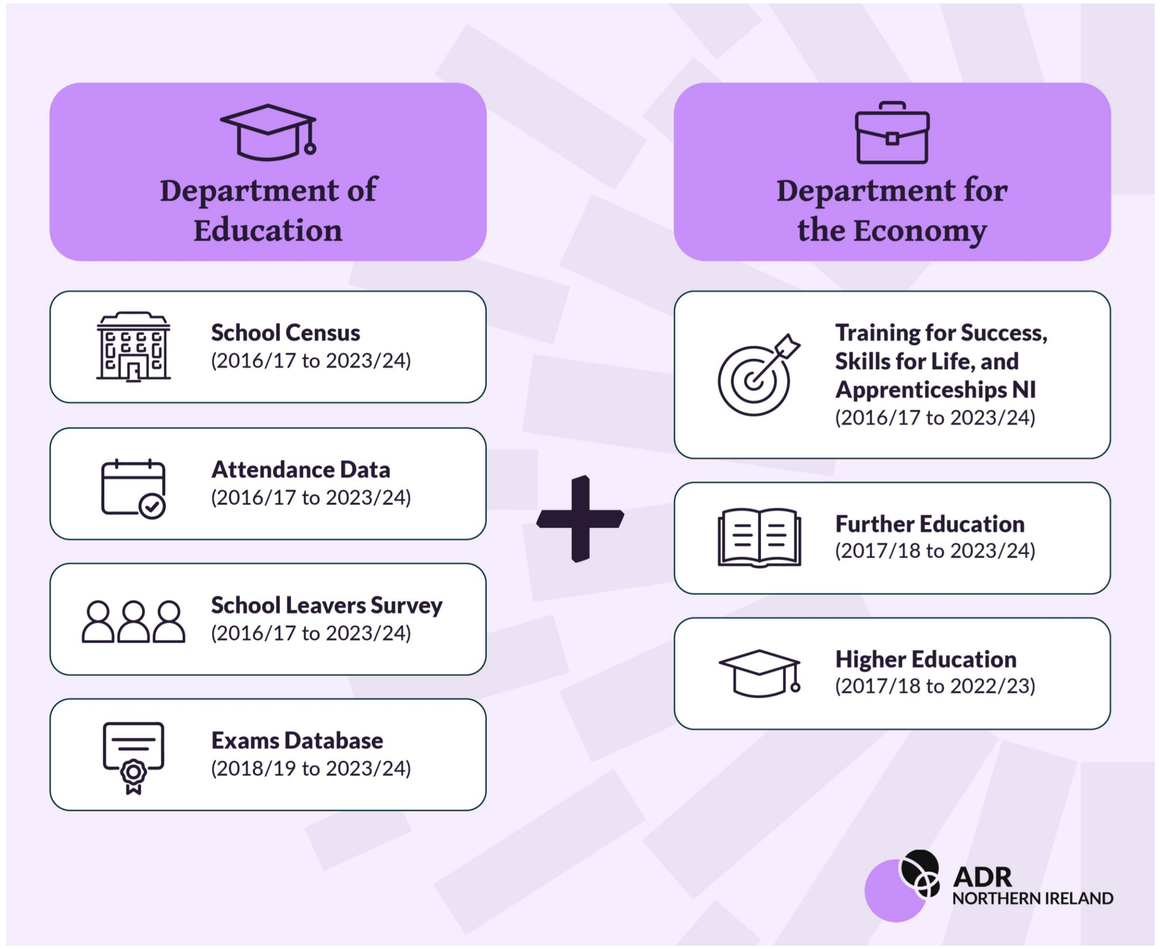
Current research uses

- Investigating health outcomes for care-experienced children
- Experiences of growing up in kinship care
- Stability of care patterns for children in Scotland

Researcher support

- **Community:** A community group for this dataset is being considered. Researchers can connect with the Children's Care and Protection Data Community.
- **Policy briefing:** This policy briefing shares researchers' experiences of using the data.
- **Technical guide:** This report documents the initial development of the dataset.

Longitudinal Education Outcomes Northern Ireland (LEO NI)



Dataset at a glance

- De-identified post-primary school census, attendance, exam and school leaver records linked with further education, higher education, training and apprenticeship data
- **Timeframe:** 2016/17–2023/24 (varies by dataset)
- **Records:** 541,790 learners (aged 14+)

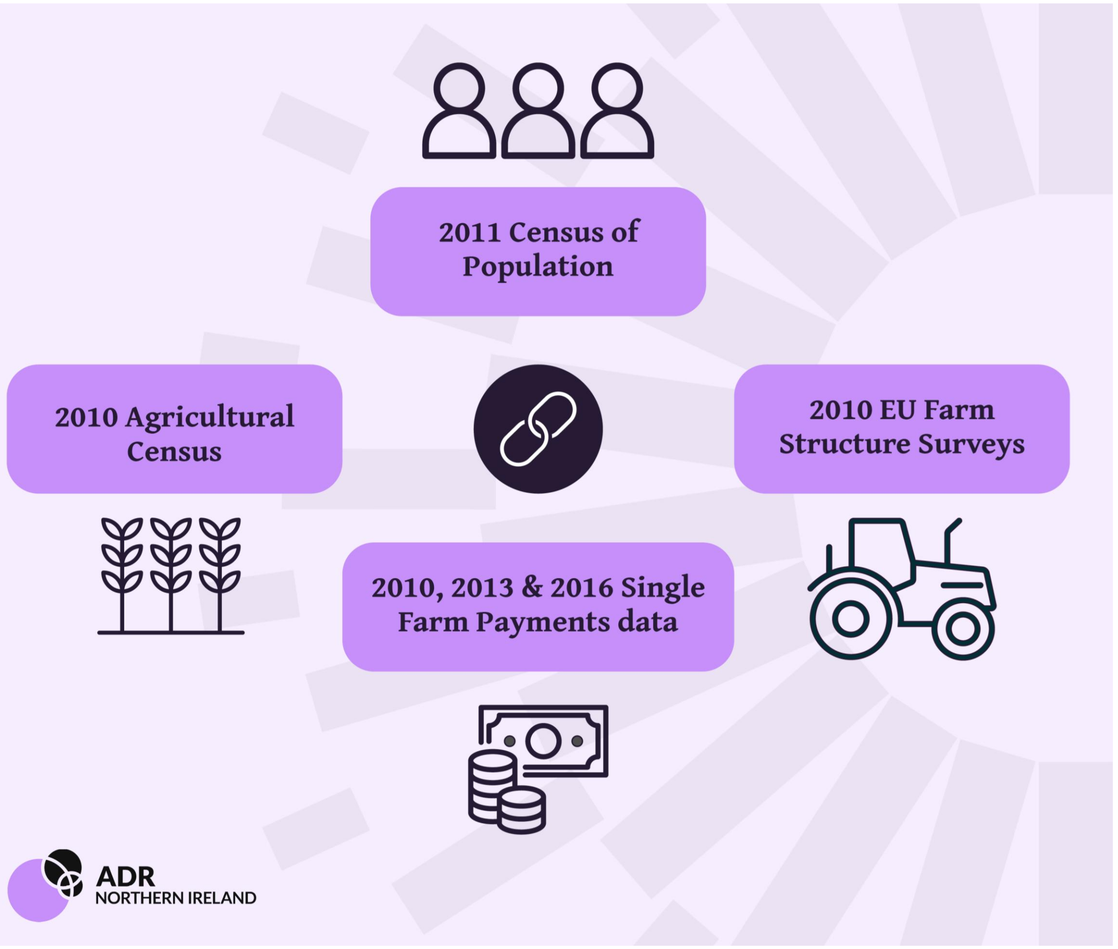
Research topics

- Impact of school attendance on academic attainment and post-16 educational progression
- Educational and progression outcomes for disadvantaged and care-experienced pupils
- Role of further education, apprenticeships and training in promoting social mobility

Dataset resources

- **Website:** Longitudinal Education Outcomes Northern Ireland (LEO NI) on NISRA
- **Synthetic data for this dataset is under development; contact rsu@nisra.gov.uk to find out more**
- **Information on the component datasets can be accessed through the ADR UK website**

Administrative Data | Agricultural Research Collection: 2011 – Northern Ireland



Dataset at a glance

- Links farm business activities, land use and subsidy payments with de-identified household census data for farming families in Northern Ireland
- **Timeframe:** 2010–2016 (varies by dataset)
- **Records:** ~74,000 individuals across ~24,000 farms

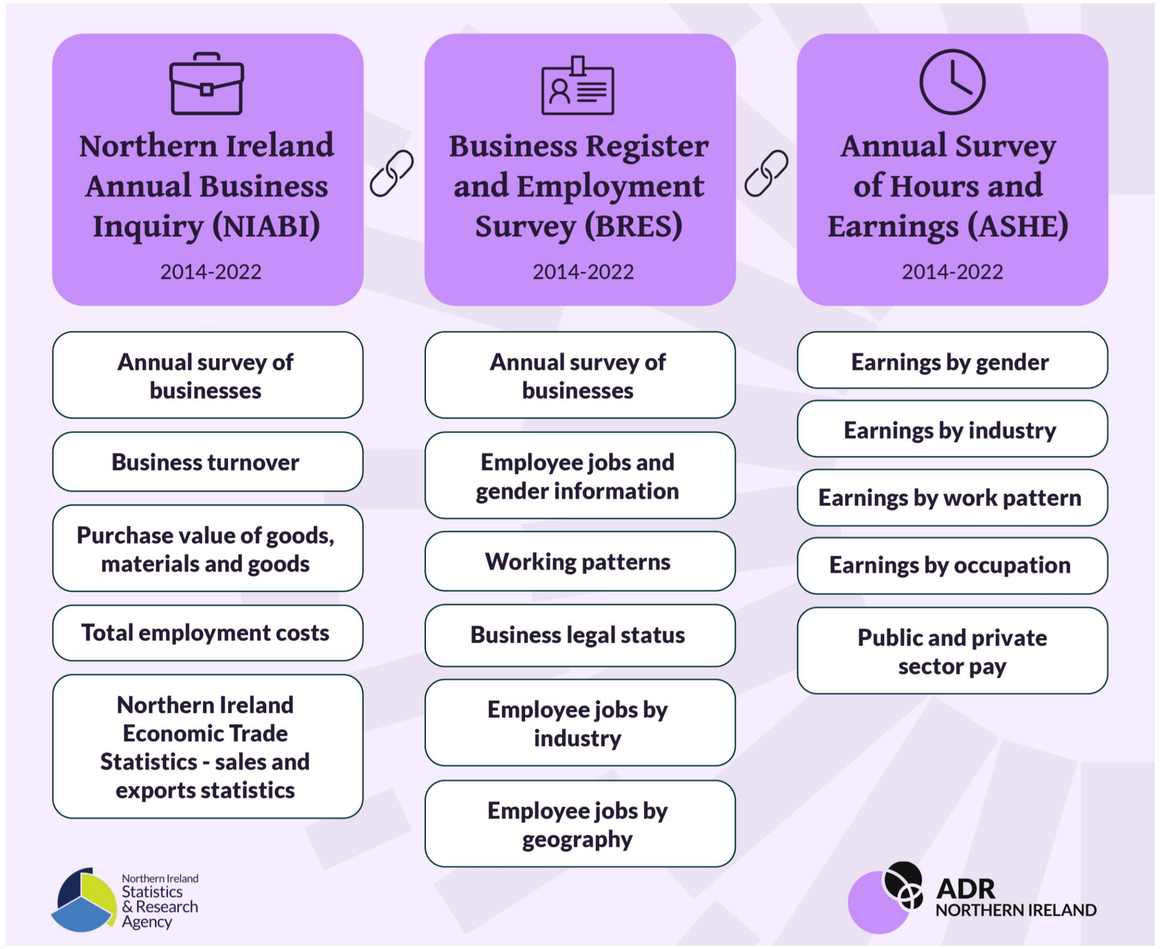
Current research uses:

- Enhancing the prosperity and wellbeing of farm households across the UK

Dataset resources:

- **User guides and metadata:** AD|ARC 2011 NI Project Documentation
- **Project website:** AD|ARC Official Website

Business Data for Research (BDR) 2022 – Northern Ireland



Dataset at a glance

- Links de-identified business turnover, employment costs, working patterns and sales/exports survey data with hours and earnings records
- **Timeframe:** 2014–2022
- **Records:** NIABI/NIETS dataset: 57,692 rows; BRES dataset: 329,657 rows; ASHE dataset: 52,931 rows

Previous research uses:

- Productivity differences across Northern Ireland local government districts
- Characteristics and employment scale of exporting versus non-exporting firms
- Impact of technology investment on employment growth, wages and productivity

Dataset resources:

- **User guide:** BDR Guidance for researchers
- **Data dictionary:** BDR Metadata and data dictionary
- **Website:** Business Data for Research (BDR) on NISRA
- **Contact:** rsu@nisra.gov.uk

Earnings and Employees Study (EES) 2011 – Northern Ireland



Links Annual Survey of Hours and Earnings (ASHE) 2011 with 2011 Census of Population and Housing

- Earnings by gender
- Earnings by work pattern
- Earnings by industry
- Earnings by occupation
- Public and private sector pay
- Population and household info
- Capital Value data

Dataset at a glance

- De-identified earnings, hours worked and property capital values linked with 2011 Census demographic, household characteristic and community background data from Northern Ireland (NI)
- **Timeframe:** 2011
- **Records:** 5,770 records (~1% of NI PAYE employees)

Current research uses

- Estimating religious community background wage differentials
- Evaluating the impact of unpaid caring on employment and earnings

Dataset resources

- **Metadata:** EES Metadata
- **Data dictionary:** EES Variable List
- **Synthetic data for this dataset is available on request; please contact** rsu@nisra.gov.uk
- **Website:** Earnings and Employees Study (EES) 2011

Dataset information for Wales is coming soon

Dataset at a glance

- Summary
- **Timeframe:**
- **Records:**

Current research uses

Dataset resources