



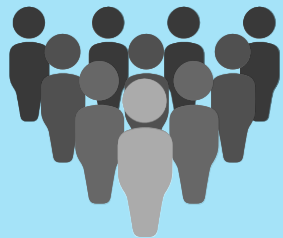
Ministry
of Justice



Research Fellowship Applicant Workshop Data First cross-justice system dataset

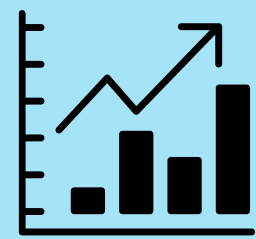
November 2025

How Data First Works



When people interact with the justice system, administrative data is created

Data First links and cleans these datasets, having developed the groundbreaking SPLINK software



These linked datasets are made available to academic researchers and other government departments via the ONS Secure Research Service and the SAIL Databank

To date, Data First has enabled over 65 research projects from government analysts and academic researchers



Dataset Overview - Data First cross-justice system dataset



Data First datasets

Criminal justice

- Ministry of Justice Data First magistrates' courts defendant
- Ministry of Justice Data First Crown Court defendant
- Ministry of Justice Data First prisoner custodial journey
- Ministry of Justice Data First probation
- Ministry of Justice Data First Offender Assessment System (OASys)

can link on
defendant and
on case

focus on the
defendant /
offender – can link
on person

cross-justice system linkage,
can link on person
(including multiple parties
in civil & family cases)

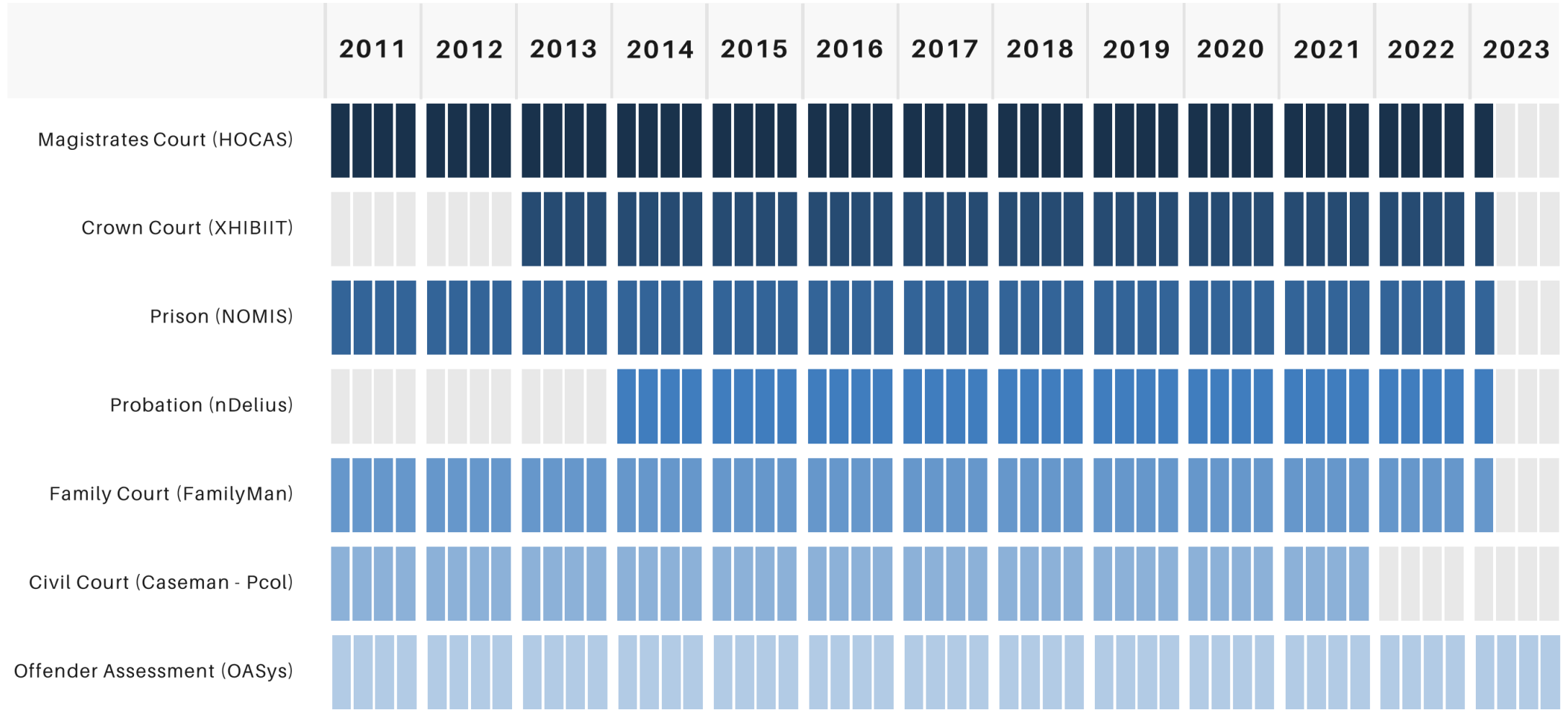
Other court jurisdictions

- Ministry of Justice Data First family court
- Ministry of Justice Data First civil court
- *Ministry of Justice Data First linking dataset(s)*

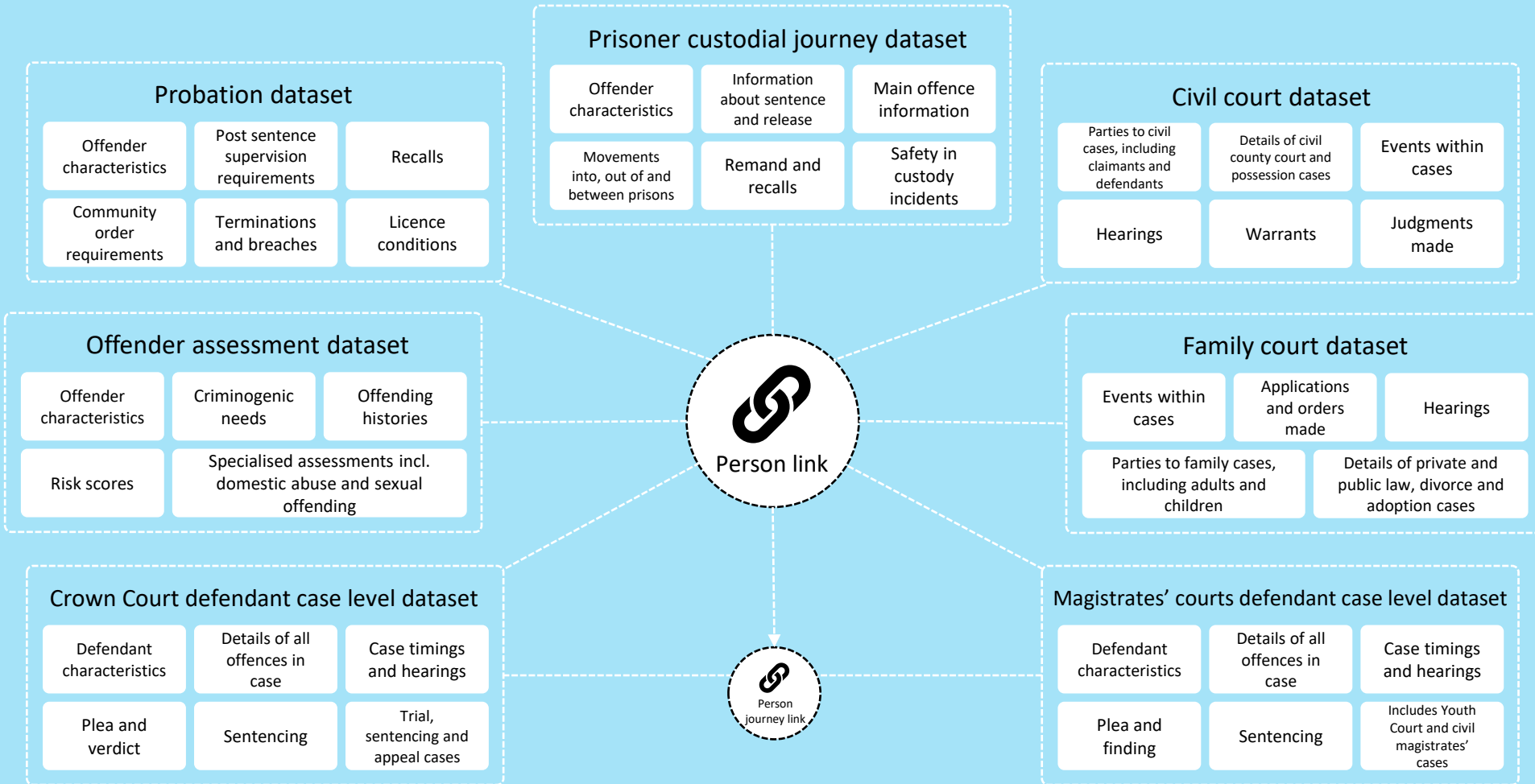
Linked to other government department's data

- MoJ-DfE data share joining the Police National Computer (PNC) to the National Pupil Database (NPD) } Separate linkage

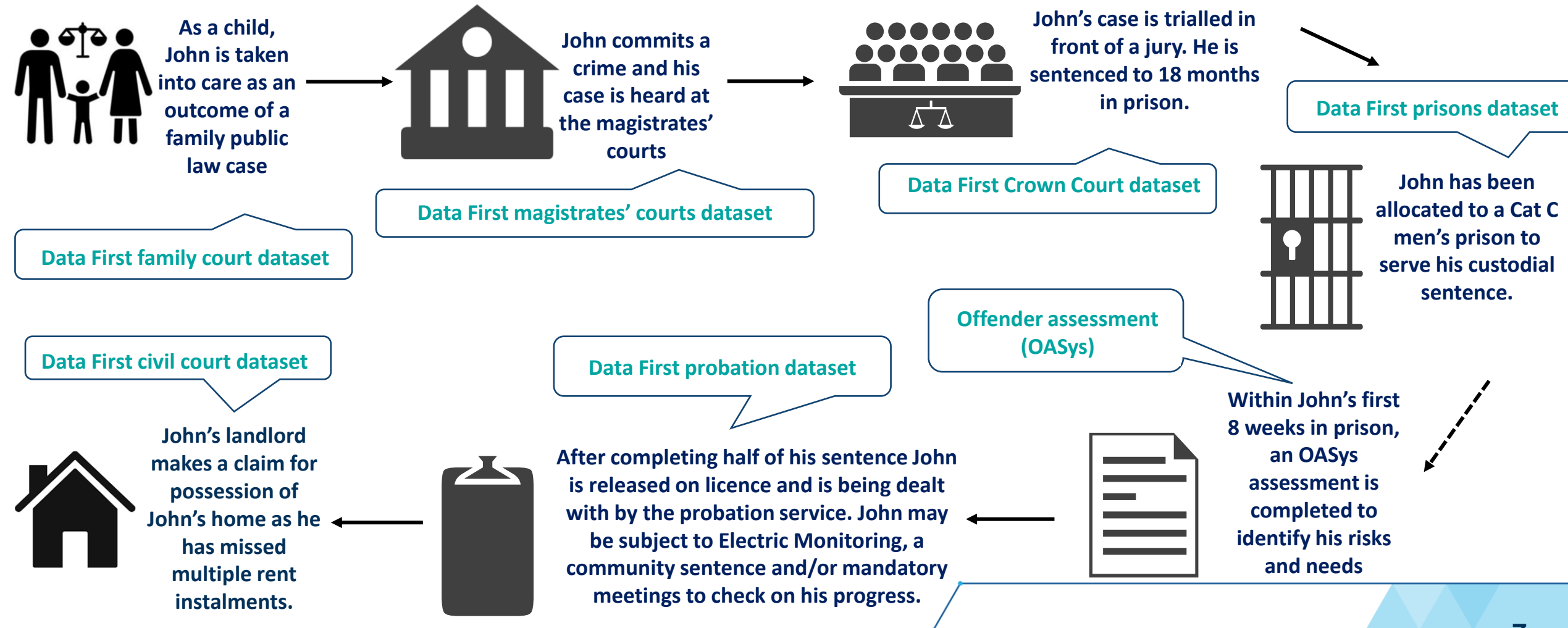
Coverage Periods



Data First cross-justice system dataset



Data First cross-justice system dataset - user journey



Previous and ongoing projects



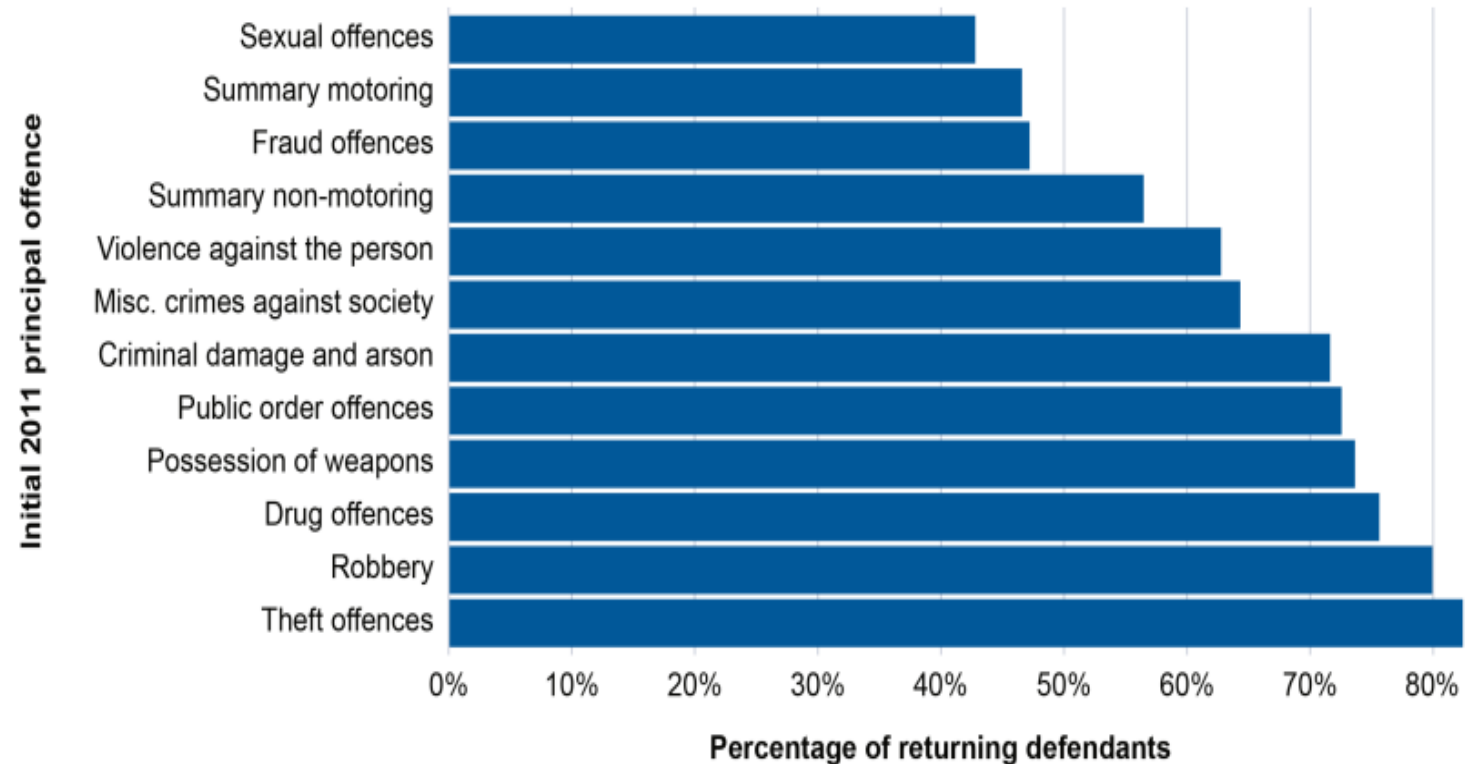
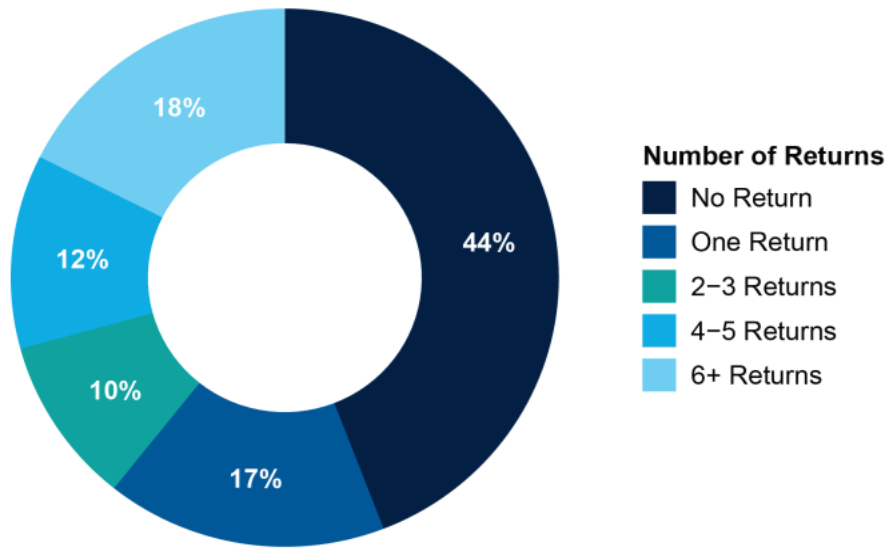
Projects utilising the Data First cross-justice system dataset

Over 50 projects have been approved to access Data First cross-justice linked datasets, with research themes including:

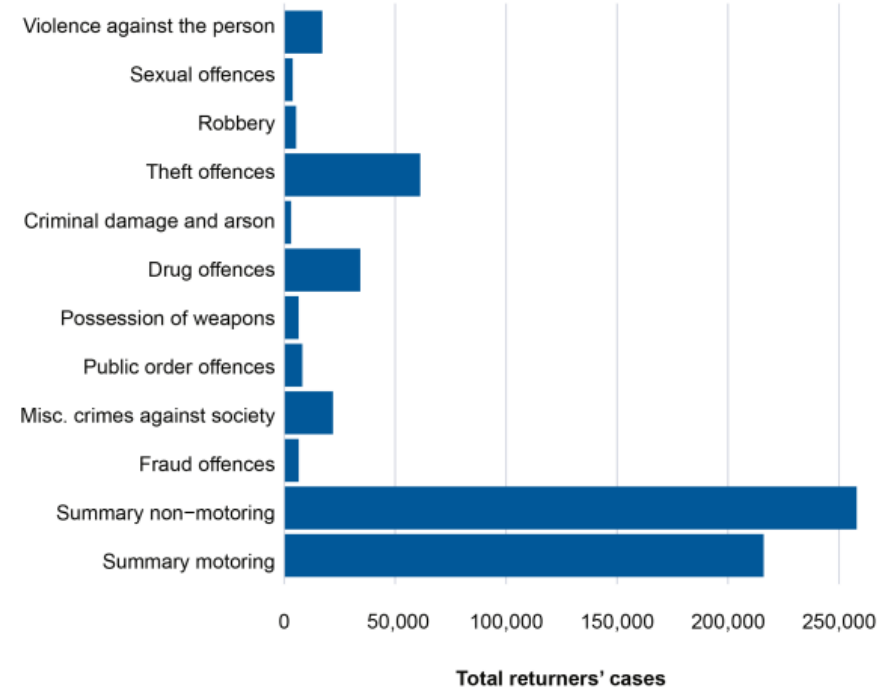
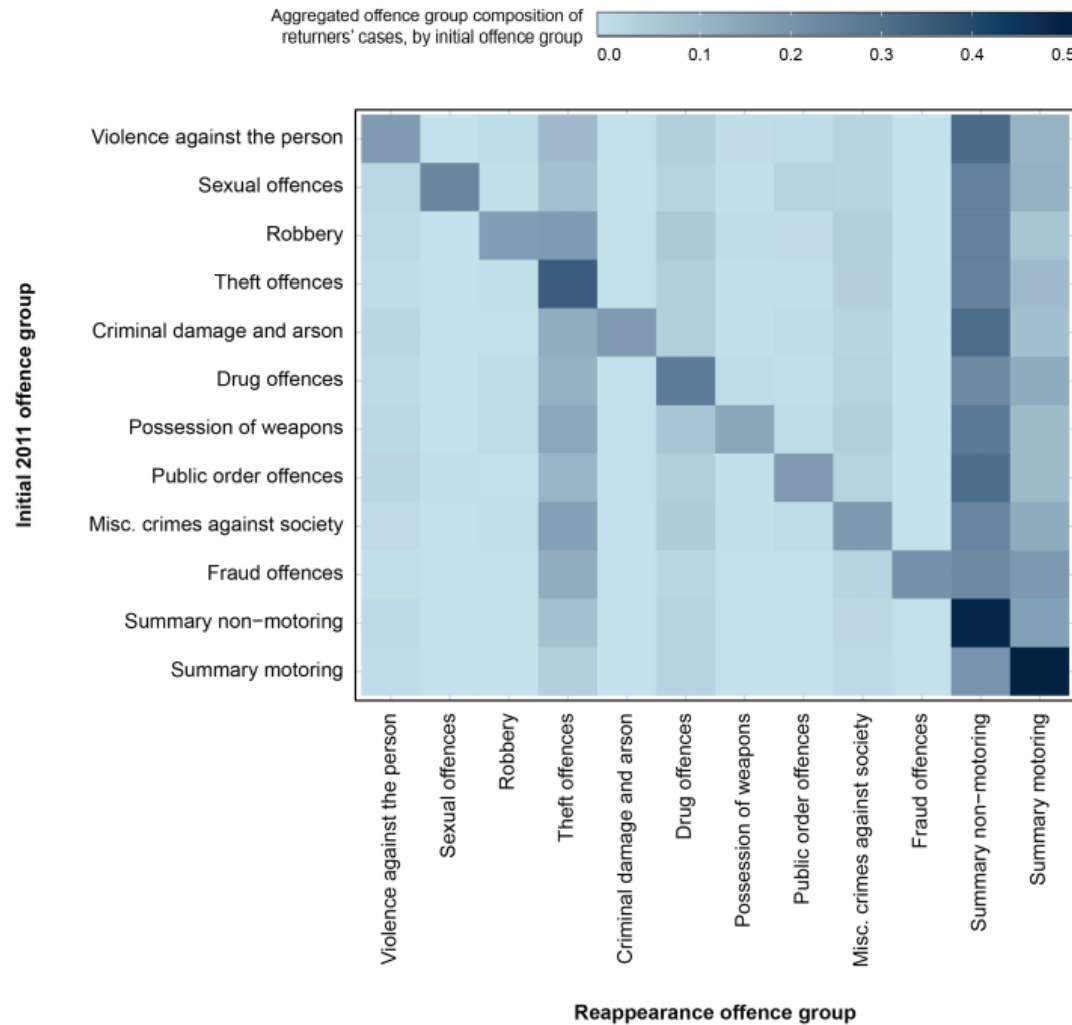
- Impact of specific policies, legislation
 - Offender Rehabilitation Act 2014
 - Implementation of Universal Credit
- Alcohol and/or substance misuse and links to crime
- Racial / ethnic differences in offence type and/or outcome
- Offence types:
 - Serious and organised crime
 - Drug trafficking
 - Sexual offences

An exploratory analysis of returning defendants and the potential of linked criminal courts data from 2011 to 2019 in England and Wales

Percentage of defendants who return to the magistrates' courts for any offence up to December 2019, by initial 2011 principal offence



Heatmap showing the proportion of defendants proceeded against by initial 2011 principal offence, to the subsequent return to the magistrates' courts



Ethnic inequalities in the criminal justice system

Research found significant ethnic disparities in sentencing outcomes after controlling for defendant, case and court characteristics

Not-guilty plea rates are higher for defendants in the Asian and Black groups, and particularly for **Black African and Indian** defendants who are 1.7 times more likely to plead not guilty than the White British group.



Among those convicted in the Crown Court, **defendants from the Chinese, Other White and Black Caribbean groups are more likely to receive a custodial sentence** than those in the White British group.




In contrast, ethnic disparities in sentence length are largely explained by legal factors, and after adjusting for these factors, observed differences between most ethnic minority groups and the White British disappear.


The exception is defendants from Pakistani, Bangladeshi and Black Caribbean groups who are shown to have worse sentencing outcomes than the White British receiving between 4% and 11% longer sentences than the White British.



- Discussions are ongoing about implications for **judicial training and guidance**.
- Research cited in the **Tackling Racial Disparities 2023** report, documenting the efforts to identify and address racial disparities in the CJS in response to the Lammy Review.

Exploring the characteristics of Prolific Young Offenders (PYO)

 Ministry of Justice

DATA FIRST
Harnessing the potential of linked administrative data for the justice system 

A Profile of Repeat Offending by Children and Young People in England and Wales
Ad hoc statistical release

Published 20 March 2025

- 27% of PYOs live in the most deprived 10% of neighbourhoods (vs. 21% of non-prolific young offenders and 11% of 18 year-olds overall)
- 27% of PYOs had at least one custodial sentence before the age of 18
- 22% of PYOs had their first convictions before age 14; 70% had their first between ages 14 and 16
- 11% of PYOs had 10 or more youth convictions before turning age 18 – suggesting a subset of particularly criminally active individuals, even within the PYO cohort

Data First research priorities



Data First Research Priorities for the cross-justice system dataset

- We welcome applications exploring policy-focused questions based on the following four themes. Other research questions are also welcomed providing you can evidence policy-relevance. Please see the [MoJ Areas of Research Interest \(2025\)](#) which sets out MoJ's evidence priorities

Cross-jurisdiction: integrations and journeys across the justice system.

- How can we better understand how problems interact and reinforce each other, and how people move through different courts across the system as they attempt to resolve them?
- Who are the 'repeat' users across the justice system? Does involvement in other areas of the justice system influence criminal (re-)offending? What works to enable effective outcomes across the justice system and reduce (re-)offending?
- How do individuals in the justice system vary and compare across the courts in the justice system? What factors drive disparities in outcomes at different stages of the court journey? How does intersectionality impact upon this?

Reducing reoffending: understanding characteristics, needs and risks factors of offenders

- How do individuals' needs change over time within the criminal justice system? What characterises individuals that stop offending?
- What are the risks and needs of different offender groups? How do these change over time and affect their likelihood of reconviction?

Data First Research Priorities for the cross-justice system dataset

Effectiveness of electronic tagging and monitoring and alcohol abstinence monitoring in protecting the public from harm.

- Are there specific groups of individuals or circumstances under which electronic tagging and monitoring are more effective? Are there specific groups of individuals or circumstances under which alcohol abstinence monitoring is more effective?
- What combination of probation requirements and conditions are most effective at improving compliance with electronic monitoring, and in reducing reoffending?

Safer Streets mission: advance our understanding and generate new knowledge to support the UK Government's objectives for the Safer Streets Mission

- What offender characteristics and needs are associated with certain types of crime (in particular: knife crime, violence against women and girls [VAWG], and anti-social behaviour [ASB]) and with reoffending across these offences?
- Is there overlap between people who have committed different types of offences, for instance between knife crime, VAWG and ASB offenders? Are there common offender trajectories across different types of offences?
- Are there patterns in how knife crime, VAWG and ASB offenders interact across the justice system?

Understanding the research potential of Data First

Administrative data can only help with certain types of questions

- Broadly speaking, Data First helps address questions about **who** is interacting with the justice system **in what ways**, the **processes and timings** involved, and **'hard' outcomes** such as sentencing or repeat appearances.
- The kind of data we have cannot provide insight into **unmet needs, circumstances and experiences** of engaging with justice services, and (at this stage) longer-term outcomes.

Data First focuses on justice system users

- Management information systems in scope for Data First focuses on **justice system users** (defendants, prisoners, parties to family or civil court cases).
- Will provide little information for questions about **the judiciary, court & prison staff and legal professionals**. Limited admin data available on some public users e.g. victims, witnesses.

A good window into the big picture and small subsets

- Data on **all** defendants, cases, offenders in custody or under supervision etc. in England & Wales.
- Can look at niche groups e.g. detailed ethnicities, specific offence or case types.

Applying to access the cross-justice system dataset



Feasibility & Resources

- Talk to us! datafirst@justice.gov.uk
 - Knowledge of datasets, completeness, pitfalls, similar projects
 - Can facilitate contact with specific data and/or policy experts
- Data catalogues, user guide, research bulletins - [Ministry of Justice: Data First - GOV.UK](#)
 - Variables lists and lookups for all datasets
 - Application process, guides to all datasets including potential research questions
 - Internal publications, previous research fellowships
- Synthetic datasets available through [UK Data Service](#)
 - Test out linking records, available variables, potential code

Applying for access to the cross-justice system dataset

Three steps for accessing data:

1. Become an [accredited researcher](#).
2. Apply for access to the specific data required for a research project from relevant data owners using the MoJ and HMCTS '[Secure access to data](#)' combined form for access to data held in the ONS SRS or complete the [scoping form and Information Governance Review Panel \(IGRP\) form](#) for data held by SAIL Databank.
3. Complete the [Research Project Application](#) form and receive final approval from UKSA RAP.

Timelines for applying to access the data

- Depends on application – including how clear the proposal, research questions and justification for specific variables are
- Feasibility, clarity, necessary documents – Data First Research, Academic Engagement & Comms team
 - Benefits to engaging with us early!
- Review panels
 - HMPPS data = Data Access Group (DAG) + Data Access Governance Board (DAGB)
 - HMCTS data = Data Access Panel (DAP)
 - Both = Both
 - Proposed methodology, data-specific knowledge, ethics, disclosure concerns
- Application can be:
 - a) Immediately approved
 - b) Approved with caveats
 - c) Returned to applicant for amendment – to be approved by correspondence, or by full panel
- Separate approval required from UKSA Research Accreditation Panel (RAP)

What makes a good application

- **Completeness**
 - **Required:** Application form, variable list **with justifications**, proof of ethical approval
 - **Helpful:** Project timeline / flow chart
- **Clarity**
 - What is the cohort your are investigating?
 - Research questions / project milestones / chapter headings etc.
 - Clear, straightforward description of methodology – clear English, limited jargon
- **Sensitivity**
 - Are you requesting any sensitive variables? If so, why?
 - Data protection concerns; e.g. low numbers?

And finally...



What you can expect from us

The Data First team are keen to support and engage throughout the end-to-end research process:

- Advice and guidance to produce a successful application for the data owner approval panels
- Provision of detailed data catalogues and synthetic data
- Availability to provide insights into dataset queries and quality issues throughout the project
- Timely review and clearance of outputs - confirm in line with approved project and advise on framing
- Informal academic peer review
- Opportunities to present and share research
- Engagement with policy

Key takeaways

- Engage with us early
- Understand the data (and allow time for exploration)
- Often, less is more!
- Clear, specific aims and research Qs
- If you have questions, please ask!



Questions?

Get in touch:

datafirst@justice.gov.uk

(if you are interested in accessing the Data First datasets, signing up to our mailing list, or for any other Data First related queries)

Find out more:

- Existing datasets, User Guide & privacy statement: <https://www.gov.uk/guidance/ministry-of-justice-data-first>
- Application form and approved projects:
<https://www.gov.uk/government/publications/moj-data-first-application-form-for-secure-access-to-data>
- Splink data linking package: <https://github.com/moj-analytical-services/splink>
- Administrative Data Research UK: <https://www.adruk.org/>
- ONS Secure Research Service:
<https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/secureresearchservice>
- SAIL Databank: <https://saildatabank.com/>