An overview of the UK’s approach to ethnic data collection in the context of the Framework Convention on the Protection of National Minorities

Craig Willis*

European Centre for Minority Issues

Abstract

In comparison with most European/EU states, the UK collects a vast amount of data disaggregated by ethnicity. This covers areas such as health, education, employment, and a variety of welfare programmes, and is not just limited to the public sector. The purpose of this article is to demonstrate how widespread this is and allow for a comparison to be made against other states’ approaches. First, an historical overview explains how the UK has come to its current position in terms of data collection, including how terminologies evolved and how such data spread from being only in the census to having a much broader scope. This scope is unpacked in detail through different areas, demonstrating the depth of data disaggregation by ethnic group in the available statistics. Lastly, an analysis of the monitoring cycles of the Framework Convention on the Protection of National Minorities offers an external perspective on the UK’s practice of ethnic data collection. The article concludes by observing that the UK continues to collect ethnicity data in an increasing number of areas and recommends that this be extended to include more minority groups.

Keywords: Ethnicity data collection, census data, Framework Convention for the Protection of National Minorities (FCNM), ACFC (Advisory Committee on the FCNM) monitoring cycles, United Kingdom (UK)

*Contact: willis@ecmi.de or cpmwillis@gmail.com
1. Introduction

The UK has a long history of data collection which has been extended to ethnic groups in recent decades and has vastly expanded in terms of the areas of society covered. Indeed, the UK was one of just three EU member states with a legal obligation to collect data on ethnicity (Farkas, 2017) – forming a Public Sector Equality Duty within the Equality Act 2010. The topic of ethnic data collection receives regular attention in the UK’s monitoring of the Framework Convention on the Protection of National Minorities 1995 (FCNM), both from the UK government and from the Advisory Committee on the FCNM (ACFC). Accordingly, this article aims to unpack the empirical practice of ethnic data collection in the UK in order to contribute to the wider European debate and offer the UK as a comparative case. To do this, an historical overview first demonstrates the legal developments around ethnicity in the progression of UK data collection throughout the 20th century and into the 21st. This then logically leads into the present-day situation, which is outlined in detail in terms of the areas of society covered and the methods of obtaining such information. This largely governmental perspective is then built on with a consequent analysis of the monitoring cycles of the FCNM and the Opinions and Resolutions developed by the ACFC and Committee of Ministers. Such analysis allows the aspect of purpose to be explored in greater detail, as well as offering an alternative to the governmental perspective.

The analysis clearly reveals that much of the ethnicity data collected and provided publicly by the UK government concerns England only – an element highlighted by the ACFC. While this article aims to provide an overview of the whole of the UK, it is beyond its scope to continuously detail the cases of Scotland, Wales and Northern Ireland. Nonetheless, an effort is made to inform the reader on where the differences lie. Moreover, Section 3 selects some of the most prominent areas in which ethnic data are collected (employment, education, health, and crime) but this is again somewhat limited. As is made clear below, the breadth of themes under which ethnicity data are available is extremely broad. Thus, the aim is very much to provide an overview and allow for consequent comparison, both within the UK context and across the wider Council of Europe area.

2. Historical overview

The UK has a long history of data collection and, almost since the start of modern census collection, there has been an aspect linked to the origin of persons. Although the first modern census in the UK was 1801, it was not until 1841 that this constituted more than just a head
count. In 1841, the census recorded the name, age, gender, and occupation of each individual, as well as where they were born. These latter questions were extremely simple, asking whether the respondent was born in the county being surveyed (England/Wales), as well as whether they were born in Scotland, Ireland, or ‘Foreign Parts’ (Office for National Statistics, n.d. a). Although a long way from ascertaining ethnicity, the 1841 census offered a first statistical look at the different nations of the UK, as well as recording the first waves of settlement of persons born outside of England and Wales, or outside the British Isles. Data from these mid-19th century censuses began to be used to analyse the population (and change) by nationality. For instance, Shannon’s 1935 analysis of London – the destination for the majority of migrants from outside the British Isles – showed that the population listed as born ‘abroad’ grew from around 43,000 in 1841 to 144,000 in 1891 (Shannon, 1935: 83); this was an example of early analysis utilizing collected data on nationality.

It was only in the post-WWII period, with its significantly larger waves of migration, that the debate around further data collection began to intensify (Owen, 1995). The UK was beginning to see a growth of a second-generation non-white population, alongside increasing numbers of commonwealth citizens coming to the UK – until 1962, there was no legislation to limit this. Thus, by the late 1950s, societal questions concerning ethnicity, race, and immigration were intensifying, and subsequently fuelled by civil rights protests against racial discrimination throughout the 1960s, these issues became politicized into mainstream debate (Bulmer, 1986). This eventually led to the Race Relations Act 1965, subsequently strengthened (given a wider applicability after criticism of the original Act) in 1968 and then replaced by the Race Relations Act 1976. Alongside this, the question of ethnicity data collection – at first mostly focused on census questions – emerged as one such politicized issue (ibid.: 472). Bulmer dissects this and outlines the development from a proxy question on race to a direct question in 1991. Until 1991, the census only featured a question on country of birth, allowing race and ethnicity to be ‘inferred’; however, Bulmer states ‘by 1971…it was recognized that an increasing proportion of the black population, the second generation, had been born in Britain and that they would not be identified by a question on country of birth’ (ibid.: 473). As a result, the 1971 census also collected information on the respondents’ parents’ country of birth (UK Data Service, n.d.), but the quest for finding a direct question became more politically pressing and the 1981 census left out the parent’s country of birth question (Owen, 1995). There were, however, ongoing experiments with terminology in sample censuses and other government surveys during these decades. For example, during the processing of the 1966 sample census data, terms
such as Old Commonwealth, New Commonwealth, and African Commonwealth were used as categories based on the country of birth question, but these were felt to be inadequate and potentially inaccurate (Laux, 2019). Furthermore, the General Household Survey, launched in 1971, required interviewers to ‘record whether respondents were White or ‘coloured’’ [single quotation marks and capitalization of white, as per the source document] (ibid.). The culmination of these trials and political debates was the decision to include a direct question on ethnicity in the 1991 census. Under the heading of ‘ethnic group’, the census form supplied nine boxes: White, Black-Caribbean, Black-African, Black-Other (with a blank write-in box below), Indian, Pakistani, Bangladeshi, Chinese and ‘Any other group’ (again with a blank write-in box below) (ibid., 2019; UK Data Service, n.d.). This style of question has become the standard for subsequent censuses as well as other surveys, such as the Labour Market Survey, which adopted this style in 1992 (Laux, 2019; Owen, 1995). Since then, the main differences in the census have been to update terminology and add further boxes, which in 2001 included the following main categories and subcategories: White (British, Irish, Any other); Mixed (White and Black Caribbean, White and Black African, White and Asian, Any other mixed background); Asian or Asian British (Indian, Pakistani, Bangladeshi, Any other Asian background); Black or Black British (Caribbean, African, Any other Black background); Chinese or other ethnic group (Chinese, Any other). The 2011 census added boxes for ‘Gypsy or Irish Traveller’ and ‘Arab’, and the 2021 version intends to include ‘Roma’ as well. Alongside these developments, the 2011 census added a separate question on ‘national identity’ which interestingly was placed before ‘ethnic group’ in the question order. This question offered the following tick boxes: English, Welsh, Scottish, Northern Irish, British, Other (write-in) (Office for National Statistics, n.d. b). Under this section, 73,200 people wrote ‘Cornish’ as their answer (Cornwall.gov.uk, 2017) – this is relevant to the instance of national minorities and the FCNM, discussed later in this article. It should also be stated that the ‘ethnic group’ question in the census in mandatory, with no option for ‘prefer to not state’ or similar; thus it is compulsory to declare one’s ethnicity.

This historical development has largely focused on the case of England (and Wales), whereas legislation in Scotland and (particularly) Northern Ireland differs. While, for the most part, Scotland has followed a similar path in terms of census terminology etc., Northern Ireland – owing to its eventful and turbulent modern history – follows vastly different legislation with significant effects on ethnicity data collection. This was relevant both before and after the Good Friday Agreement in 1998; the Race Relations Act 1965 (and 1968 and 1976) of the UK did
not apply to Northern Ireland, and neither does the current *Equality Act 2010*. Instead, Northern Ireland’s core legislation in this field is the *Race Relations Order (Northern Ireland) 1997*, amended in 2003 and 2009 (Equality Commission for Northern Ireland, 2011). This Order was put in place alongside the Good Friday Agreement which stipulates that the people of Northern Ireland have a ‘birthright…to identify themselves and be accepted as Irish or British, or both, as they may choose’ and that ‘British and Irish citizenship is accepted by both Governments’ (Belfast Agreement, 1998). In terms of ethnic data collection, this has been interpreted to mean, in practice, that it is not lawful to ask people in Northern Ireland to choose between identifying as Irish or British; this contrasts with censuses of England and Wales, as well as Scotland (Office for National Statistics, n.d. c). As such, in the 2011 Northern Ireland census, the ‘ethnic group’ question offered the following categories: White, Chinese, Irish Traveller, Indian, Pakistani, Bangladeshi, Black Caribbean, Black African, Black Other, Mixed ethnic group (write-in), Any other ethnic group (write-in). There is no mention of the term ‘British’ whatsoever, and no subcategories as with the England and Wales census. In the question on ‘national identity’, it is the same as England and Wales in that it stipulates ‘please tick all that apply’ (UK Data Service, n.d.).

The present-day situation is thus a consequence of these historical developments, whereby the format used in each census is generally the blueprint for all other ethnicity data collection. Furthermore, the practice is now much more widespread than the snapshot provided by the census every 10 years. Ethnicity data are now collected in relation to education, welfare benefits, employment, the national health service (NHS), the prison service, and more. The following section unpacks these various data collection areas in more detail and summarizes the stated purpose and rationale given for the ethnicity data collection in each case.

3. Areas of ethnic data collection and its purpose

Given the widespread fields in which ethnicity data are now collected in the UK, and the varying justifications provided by UK governmental departments for doing so, it is useful to break this down into subcategories. Each contains a summary of how ethnicity data are collected – most are collected through relevant population surveys or more directly through administrative processes. It should be added that, generally, the UK authorities are very vague on their justifications for collecting ethnicity data – the general guidelines provide no mention of this whatsoever (Office for National Statistics, n.d. c). However, there is some information on the stated purpose in a few areas; this is detailed within the relevant subsection below.
To this end, the *Equality Act 2010* contains a Public Sector Equality Duty (PSED) which sets out requirements for public institutions to adhere to (this replaced previous separate duties in the legislation superseded by the *Equality Act 2010*).\(^{10}\) This has relevance in a number of areas, as will be detailed in each section, but in general refers to three key aspects:

- Eliminate unlawful discrimination, harassment and victimisation and other conduct prohibited by the Act.
- Advance equality of opportunity between people who share a protected characteristic and those who do not.
- Foster good relations between people who share a protected characteristic and those who do not.

(Equality and Human Rights Commission, n. d.).

As part of implementing the PSED, public authorities\(^ {11}\) are obliged to collect and publish data on employees and ‘people affected by its policies and practices (for example, service users)’ (Equality and Human Rights Commission, 2015: 5). This includes data on ethnic group, as well as other protected characteristics covered by the Equality Act. Thus, the general purpose is stated to be that such data are necessary to demonstrate public authorities’ compliance with the duty.

In addition to the public authorities’ requirements, the UK government provides general guidance for those wishing to collect data on ethnicity – these include public authorities, private employers, survey administrators, and individual researchers. The guidance recommends using its harmonized ethnic group question and categories, which it states will ‘enable consistency and comparability of data’ (Office for National Statistics, n.d. c). The guidance suggests opening with the question ‘what is your ethnic group’ and providing categories as used in the census. Although it acknowledges that the categories may impact the individual’s choice and limit their ability to self-identify, it justifies this with the need to organize data for useful comparison (ibid.).

As a consequence, ethnicity data are aggregated and made available to the general public through a government website categorized under eight broad topics: crime, justice and the law; culture and community; education, skills and training; health; housing; UK population; work, pay and benefits; and workforce and business (Gov.uk, n.d. a). Additionally, there are in the region of 100 subcategories which are very wide reaching. This Gov.uk portal mostly contains data concerning England only; however, there are separate websites for Scotland, Wales and Northern Ireland which focus on equality data including ethnicity.\(^ {12}\) For the purposes of scope and to avoid repetition, this article focuses mostly on the data for England, or UK wide where
possible. Similarly, it is not possible to cover all 100-plus areas in which ethnicity data are available. As such, the following subsections concentrate on some of the core areas and provide examples therein.

The Data Protection Act 2018 is the UK’s current legislation relevant in this area, incorporating the EU’s General Data Protection Regulation (GDPR). This British legislation thus follows the EU regulation on processing sensitive data, which includes ‘personal data revealing racial or ethnic origin’, among other aspects (Data Protection Act 2018, Paras. 35.8 and 86.7). This is detailed further in Schedule 1 of the legislation, which covers data protection in the context of ‘equality of opportunity or treatment’ and ‘racial and ethnic diversity at senior levels of organisations’. The government guidance on the Data Protection Act 2018 further outlines the general areas of responsibility regarding personal information and how it is used by government, organizations, or businesses. Here it states that ‘there is a stronger legal protection’ for sensitive information such as race and ethnic background, and outlines the rights individuals have – although the terms are quite general and do not specify the method through which these can be realized (Gov.uk, n.d. b). Certain government departments, such as the Department for Education, offer more specific guidance (Gov.uk, 2019).

3.1 Employment

When discussing ethnic data collection in the context of employment, there are two differentiations to be drawn. First, administrative data are distinct from survey data; second, administrative data can be subdivided into private– and public sector employment. Public sector employment is covered by the PSED outlined above, and as a consequence, it is compulsory for public authorities to collect data on their employees. In the private sector, this is optional, but it is important to highlight that it is legal for companies to do so, providing that they comply with the legal regulations, including data protection law. As outlined above, the UK government guidance suggests using the harmonized ‘ethnic group’ question so that data are easily comparable between sectors. Moreover, the purpose (at least for the public sector) is demonstrated through the aspects outlined in the PSED for which the data allows the duties to be fulfilled. On top of the data and analysis made available by the UK government (which is outlined in the paragraphs below), there are also academics using this data in research to highlight inequalities and recommend policy – thus offering further evidence of the purpose of such data collection. The following examples of research topics are by no means exhaustive, but demonstrate the breadth of topics covered: under-representation of ethnic minorities in the
British armed services (Mason & Dandeker, 2009); over-education and consequent under-skilled usage of non-white groups in the labour force (Lindley, 2007); and the effects of Universal Credit and lower wages of ethnic minorities (Clark & Shankley, 2020).

In terms of how ethnic group data related to employment in the UK is collected, administrative data are the most comprehensive for the public sector, given that authorities with over 150 employees are legally required under the PSED to collect such data. As a consequence, there are publicly available data broken down by ethnic group, age, gender, employment grade, and also across time (Gov.uk, 2018a). This includes a variety of simple graphs and tables integrated into the Gov.uk webpage in a generally very accessible manner without any need for downloading excel spreadsheets or self-selecting variables – as can often be the case with government statistics.

In terms of the private sector/wider general public, macro employment statistics are available, broken down by ethnic group as well as by gender and age group (Gov.uk, 2018b). These comparisons are then also available across time, with data available from 2004 onwards. The data source for this is the Annual Population Survey, which also uses regional data from the Labour Force Survey. The Annual Population Survey uses a sample size of around 320,000 respondents and is stated to be a ‘continuous household survey’, in operation since 2004 and collected quarterly (Office for National Statistics, n.d. d). What this means in practice is that it is a representative sample of the UK population; it uses systematic sampling by postcode and NHS communal accommodation data, and applies weighting factors among others to ensure accuracy.\(^\text{13}\) Similar data on ethnic group is also available for unemployment, economic inactivity, pay and income, and benefit payments (Gov.uk, n.d. c). These aspects overlap with the field of employment, although for reasons of scope and to avoid repetition on methodology, not all of these aspects are specified in detail here. To touch on one aspect, welfare in relation to supplementing or replacing labour income is also subject to ethnic data collection. Disaggregated data can be viewed on Jobseekers’ Allowance in terms of total numbers, average duration of claim, and sanctions (where benefits are stopped or reduced). The data are obtained through administrative sources meaning, in practice, that Jobseekers Allowance forms include a mandatory question on ethnic group within the personal details section; however, the claimant can choose to tick a box that says ‘I do not wish to state my ethnic group’ – an option which is not available in the census.\(^\text{14}\)
3.2 Education

A wide range of data is collected on pupils in the UK throughout their time in the education system, from pre-schoolers to post-graduate university students. This involves government departments, local authorities, schools, colleges, and universities collecting data on ‘pupil attainment, absences and exclusions […] and where people go after leaving education’ (Gov.uk, n.d. d). The majority of these data come from the National Pupil Database of the Department of Education, which has access to the personal data of all pupils in the state-run system. Under this database, each child is assigned a unique number which stays with them throughout their compulsory education (Department for Education, 2015). Attached to each pupil’s data profile is a variety of personal information including ethnic group, which can then be obtained by the Department of Education at any point in an aggregated form. Pupils’ ethnic group is ascertained through a question in the personal information form filled in by parents when the student initially starts school (usually around ages 4 to 5). As with employers, schools are recommended to use the harmonized ethnic group question and categories; however, it appears that schools/districts use their own forms rather than a UK-wide form. Furthermore, at the beginning of each school year, parents (or students, once older) are given the chance to review their personal information at the beginning of each school year – although this is more relevant for change of address etc. and it seems unlikely that ethnic group would change. Ethnic group is also included on the school census which schools are required to submit each term. The scope of the school census was expanded in 2016 through new legislation, requiring schools to also collect data on nationality and country of birth (Department for Education, 2016). However, this was extremely controversial and media stories emerged of schools being told to guess pupils’ ethnicity if it was left blank on the form (Pells, 2016), or schools demanding to see pupils’ passports to check their legal status (Whittaker & Camden, 2016). In the face of strong civil society campaigning, the UK government dropped this requirement after just two years, instructing schools that they no longer needed to collect data on country of birth or nationality of students (DoE, 2019; Whittaker, 2018). The ethnic group question remained in the school census and the collection of this was largely absent from the debate, which focused more on discrimination against migrants and the government’s intention to use such data for Home Office activities (Busby, 2018; Whittaker, 2018).

As a result of collecting ethnic data, the Department for Education, and consequently the general public, have access to a wide range of data on school pupils which can be disaggregated by ethnic group. Figures on academic achievement can then be obtained and broken down by
ethnic group as well as geographical area. Data on destinations after education are also available by ethnic group, as are data on school absences and pupil exclusions (Gov.uk, n.d.d). A final note to add, as per above, is that data on teachers’ ethnic group can also be obtained due to the public sector employment requirements on data collection – again with the possibility to disaggregate them by area, gender, type of role, and type of school (Gov.uk, 2018c).

Aside from the general points made in the PSED, the Department for Education (2015) provides specific information on the purpose of collecting ethnic group data in the education sector. It states that such data are useful for schools, local authorities, and the Department for Education (at the national level) to: ‘Identify barriers to achievement; Establish strategies to raise standards; Comply with equal opportunities legislation and the Race Relations Amendment Act; Ensure effective allocation and targeting of funding’ (ibid.: 28). Furthermore, there is evidence that academic research has been undertaken using ethnic group data to analyse and highlight inequalities, as well as to make recommendations for future policy (Croxford & Raffe, 2014; Strand, 2015; Swann & Law, 2010).

3.3 Health

Under the category of health, the Gov.uk website provides 37 outcomes listed under six subcategories, all which can be disaggregated by ethnicity. The six subcategories are: access to treatment; exercise and activity; patient experiences; patient outcomes; physical and mental health; and preventing illness (Gov.uk, n.d.e). These data provide a vast variety of indicators, broken down by ethnicity, involving both the health attributes of patients and their experience within the NHS; in some cases, these are viewable across time, gender, age etc. To provide a sample of this breadth, ethnicity data are publicly available on: detentions under the Mental Health Act; physical activity and inactivity; early cancer diagnosis; adults reporting suicidal thoughts, attempts and self-harm; illicit drug use among adults; cigarette smoking among adults; and many more. It should be pointed out that many of these data are based on surveys using representative samples, and so not necessarily exact data for the country as a whole. Moreover, the ethnicity data are only available for England in most cases.

These data are collected from multiple surveys, varying in terms of frequency, size, and place of collection. This paragraph details a few of the most prominent. One of the largest, in terms of sample size, is the annual GP Patient Survey which falls under the NHS Outcomes Framework. The GP Patient Survey is sent out to around 2 million households in England each
year through the postal system, and, as an example, the 2017 figure for returned and completed questionnaires was around 808,000 (NHS England, 2017). This makes it one of the largest population surveys in the UK, and ‘what is your ethnic group’ is one of the questions under the section ‘Some questions about you’. As per the government guidelines, the form uses the harmonized ethnic group question and categories, also allowing for individuals to write in their own answer under one of the five write in boxes. This allows the entire survey (62 questions in total) to be disaggregated by ethnic group and thus is one of the major sources of government data on health and ethnic group. Aside from the Gov.uk website and the NHS statistics webpages, there is also a standalone website which focuses solely on the GP Patient Survey and allows users to search for specific GP surgeries in England to see how they fare in terms of patient experience. However, this focuses on general experience at the macro level and does not provide any data disaggregated by demographics such as ethnicity, gender, or age. Finally, it should be pointed out that this survey is just for England – there are separate similar examples for Wales, Scotland and Northern Ireland for which outcomes can be viewed separately on the respective NHS/local government webpages.

The NHS Outcomes Framework provides public information online outlining goals for the NHS in terms of outcomes and states the reasons why data are needed for this purpose (NHS Digital, n.d.; NHS Digital, 2018). There is no specific mention in the framework as to why ethnic group is needed for such data. However, at the local level, an official leaflet from the Wrightington, Wigan and Leigh NHS Foundation Trust (2011) outlines three core reasons why the NHS collects ethnic group data. These can be summarized as: understanding the needs of patients (to provide better and appropriate services); identifying patients at risk; and complying with the law (to promote race equality). These reasons are also supported by academic literature which reiterates the need for ethnic data to assess inequalities and improve outcomes, as well as to determine high-risk groups (Iqbal et al., 2009). Furthermore, ethnicity data are also already being used in academic research on health (Mathur, Grundy & Smeeth, 2013; Smart & Harrison, 2017). Thus, it appears that ethnic data collection is already influencing healthcare policy and decision-making. However, such data collection is not without controversy; the media have broken stories which highlight related issues/scandals and academics have flagged issues of definition and classification. Research published by the Royal Statistical Society in 2014 questioned whether NHS ethnicity data truly reflect patient ethnicity (Saunders, 2014; Siddique, 2014). Researcher Dr Katie Saunders found that accuracy of ethnicity was much higher for White British than for ethnic minorities, where there were mistakes up to 40% of the
time. Although no concrete cause for this could be given, the author suggested that instances of ‘guesswork’ on the part of relatives or hospital and care staff could be a significant factor in cases where a patient did not or could not self-identify their ethnicity. The major problem highlighted by the report is that, given that the NHS uses ethnicity data to identify patterns and at-risk groups, inaccurate data could be detrimental to this aim. Similarly, Mathur, Grundy & Smeeth (2013) point out that ethnicity classification can change over time – both from an individual’s personal perception and in terms of categorization. Thus, over-reliance on time-series data may also lead to inaccurate outcomes.

3.4 Crime, justice and the law

This final subsection concerns crime, justice, and the law. This is a particularly salient area given the history of the relationship between ethnicity and crime, as highlighted by the Stephen Lawrence Inquiry and the consequent recommendations and public awareness. The relevant section on the Gov.uk webpage lists 21 outcomes under the following four subcategories: policing; crime and reoffending; courts, sentencing and tribunals; prison and custody incidents (Gov.uk, n.d. f). This allows the reader to view statistics on a broad array of areas, including (to name just a few): stop and search; arrests; confidence in the local police; victims of crime; average length of custodial sentences; young people in custody; violence involving prisoners; and use of force on young people in custody. All of these have data disaggregated by ethnic group, often comparable across time and other demographics such as gender, as well as by local area in some cases. The list of statistics demonstrates a two-fold nature of the ethnic group data with relation to perpetrators, actions by and against. That is, it is possible to assess the actions of the perpetrator (including victims) as well as the actions by the institutions (police, courts, prisons) against the perpetrator. This demonstrates that the intention of the data collection is to support policies to protect the general public from crime (through theoretically more effective policing) and also how the data collection can be used to protect perpetrators (or the accused) from the institutions – this is in line with recommendations in the Stephen Lawrence Inquiry which call for ethnic data to be recorded (Macpherson, 1999: 381).

In terms of sources of data collection for crime, this is again a mix of administrative and survey data although a greater proportion of the former in comparison to the subcategories outlined above. Administrative data comes from various sources within the public sector, related to the area of relevance to the statistics. For example, ethnic data on sentences and custody is collected through the Ministry of Justice into a collection of data entitled ‘Criminal Justice
Statistics’ which is collected in a mix of annually and quarterly (Gov.uk, 2015). Other administrative sources include: Police powers and procedures England and Wales statistics collection; Race and the criminal justice system 2016 national statistics; youth justice statistics (Gov.uk, n.d. f). The author was unable to find information detailing the collection procedure of such administrative data; thus it is unknown whether this is through self-identification or not. Regarding survey data, the ‘Crime Survey for England and Wales’ (CSEW) is one of the largest used, collected since 1982 every two years and since 2001 annually (Office for National Statistics, 2019). This is a face-to-face survey designed to be ‘representative of the population of households in England and Wales’ with a sample size of 35,000 and a minimum of 650 interview in each police force area and using the postal address file for sampling (ibid.). Interestingly, these sample size figures were reduced in 2012 from 46,000 and 1,000 interviews in each police force area, for which the ONS website states was the outcome of the Government’s Comprehensive Spending Review (ibid.) – put simply, because of government spending cuts. Furthermore, the previous methodology was criticized and consequently changed from 2018 following an independent review. The ONS website specifies that CSEW respondents are ‘asked to make a choice from a card to identify their ethnic background using a standard National Statistics harmonized classification’ (ibid.) – that is, the system outlined at the start of Section 2 of this article.

Alongside the purposes stated by the UK government, there have also been academic researchers utilizing ethnic group data on crime in order to identify inequalities and recommend future policy changes. These largely use the survey data from the CSEW/BSC, given the length of data available back until 1982, and focus on criminal acts as well as judicial processes (Smith, 2005; Kautt & Tankebe, 2011). This offers some evidence of how ethnic data are being utilized beyond strictly government usage, similar to the three areas previously discussed.

3.5 Summary of core areas

The four subsections mentioned above offer in-depth examples of where and how the UK collects ethnic data, but this list is certainly not exhaustive. The extensive administrative and survey data in place ensures that ethnic group data can be found across even further reaching societal areas, also encompassing the areas of housing, community, and even cultural aspects such as museum visits. Although these four subsections only touched on a few indicators, they suffice to illustrate the methodology, sources, and procedures for gathering such data.
Therefore, the reader should now have a reasonable comprehension of how the UK collects ethnic group data and how widespread this practice is.

It appears that ethnic data collection in the UK largely runs along the lines of ‘the more the better’, as the extremely wide array of subject areas demonstrates. As the research on inaccuracies in NHS data suggests, the UK authorities appear keen to obtain as much data as possible under the presumption that it will be useful. The purpose of collecting ethnic data was outlined to a certain extent in some areas but was certainly vague in places. However, academic use of ethnic data in research, or the Stephen Lawrence Inquiry’s call for police to record ethnic data in activities such as ‘stop and search’, offer examples to the question of purpose.

Yet, despite the UK government’s direct assertion that such data help to set policies, comply with laws, and increase equality of opportunity, oftentimes little detail was provided, nor specific actions taken. Nor is it clear whether such data collection is improving the situation of ethnic minority groups or whether it just allows the issues to be highlighted. This is particularly the case for education (and crime to a lesser extent), with academic research using ethnic group data to highlight vast inequalities that exist between ethnic groups in terms of educational attainment, access, and discrimination. Healthcare gave the strongest justification for ethnic data collection and it appears that data are actually being used to inform decisions. Nonetheless, the UK government’s official correspondence remains quite muted and vague on the outcomes that ethnic group data collection can produce. Thus, it is useful to explore an alternative source of information, where the UK government is obliged by a non-governmental actor to explain and provide a commentary on its actions. For this, the monitoring cycles of the UK under the FCNM are analysed.

4. **FCNM monitoring cycles**

The focus of this article now turns to national minorities in particular and how ethnicity data are perceived. The aim here is to examine an external actor’s view of the UK’s situation, as opposed to the justifications given by the UK itself. For this, the *Opinions* of the ACFC and the *Resolutions* adopted by the Committee of Ministers are analysed for instances where data collection is mentioned. Since the UK’s signature of the FCNM in 1995 and subsequent ratification in 1998, four monitoring cycles have been completed to date. The cycle begins with the FCNM signatory submitting a *State Report*; the ACFC provides *Opinions* on the report, the signatory responds with *Government Comments* on the ACFC *Opinion*, and a *Resolution* is
ultimately adopted to end the cycle– thus the cycle comprises four documents. This section examines these chronologically from the first cycle onwards.

First, a short summary of the ACFC’s general stance towards data collection. In one of its thematic commentaries, the ACFC includes a subsection on data collection in relation to self-identification, stressing that this notion is of upmost importance (Advisory Committee on the Framework Convention for the Protection of National Minorities, 2016: 8–9). Furthermore, under a subsection on equality, the ACFC provides a clear statement of its consideration of ethnic data collection – worth quoting at length:

In this context of special and targeted measures for the promotion of effective equality, the Advisory Committee has consistently emphasised the importance of regularly collecting reliable and disaggregated equality data related to the number and situation of persons belonging to national minorities. It has, however, cautioned states parties against the over-reliance on statistics and encouraged the authorities also to avail themselves of independent research, in particular when carried out by persons belonging to national minorities themselves, in order to assess and comprehensively address the particular shortcomings faced by persons belonging to national minorities (see also paragraph 18) (ibid.: 25–26).

The first sentence unequivocally shows support for data collection on national minorities, highlighting that it has ‘consistently emphasised’ this stance, implying continuity. Its second sentence then stresses further that states should not rely on independent researchers or national minority actors to collect such data, strongly hinting at government or non-governmental action. In the context of the UK, data collection is undertaken by the State (as discussed) or by a company such as Ipsos Mori on behalf of the State – it would seem the ACFC in their general guidelines are highlighting a process which the UK is already undertaking. Analysis of the four documents in each of the four monitoring cycles will allow for a concrete understanding of this.

Regarding the definition of an ethnic or national minority, this is of course a heavily debated topic, and terminology varies greatly from state to state across the Council of Europe members. While the ACFC seeks not to exclude any minority populations and indeed also begins to include new migration populations in its Opinions, the extent to which states accept or engage with this varies. Moreover, numerous scholars and minority experts have continued to question whether the FCNM should be so limited and to what extent the ACFC should push for widening this in its Opinions (Marsal, n.d.; Medda-Windischer, Boulter and Malloy, 2019). It is therefore the case that the FCNM and the ACFC treat the definition and scope of national minority
differently state by state. In the case of the UK, the government’s approach was already clear in the introduction to the first State Report in 1999; in Part I (2.) the UK government states that “‘national minority’ is not a legally defined term within the UK’ (United Kingdom, 1999: 4). It goes on to clarify that the application of the UK’s ratification of the FCNM is based on the definitions in the Race Relations Act 1976 which covers racial groups as well as ethnic minorities, such as those with national origins within the UK (Scots, Irish, and Welsh – extended in 2014 to include Cornish), and that this is not limited to ‘race’. Therefore, it is obvious that the minorities falling under the scope of the FCNM for the UK’s implementation are broad. This is of course relevant to ethnic data collection processes in the UK where data covers groups which could be divided into old and new minorities. This contrasts with other state contexts where data might only be collected on migrant populations and not ‘national / autochthonous’ minorities, which is not the case in the UK.

4.1 First cycle

In the first cycle, the UK submitted its State Report with several mentions of the ethnic data collection and its relevance. Bearing in mind that as this was submitted in 1999, this pre-dated both the Equality Act 2010 and the PSED that was included in that substantial legislative overhaul. As such, the report refers to various pieces of legislation which required governmental departments and agencies to ‘monitor data on ethnic origin’ (United Kingdom, 1999: 49). This was summarized under a subsection entitled ‘ethnic monitoring’ where the UK also submitted statistics on ethnic minority staff in the UK civil service. The Opinion response from the ACFC only included brief references to data collection, stating that it ‘welcomes the substantial effort made by the United Kingdom authorities in preparing the State Report and including valuable data and information not only on legislation but also on relevant practice’ (Advisory Committee on the Framework Convention for the Protection of National Minorities, 2002: 4) – a clear reference to the existence of ethnic data collection and demonstrating approval of such. In its Government Comments, the UK made several more references to data collection in response to specific points in the Opinion. This was mostly to state that forthcoming data on ethnicity would be one of the UK’s core tools for addressing outstanding issues. For example, in Article 12, the ACFC expresses concern about lower attainment levels of certain groups of ethnic minority pupils; the UK responds that annual data collection ‘will enable us for the first time to monitor on a consistent basis the different levels of performance among different ethnic groups of pupils’ (United Kingdom, 2002: 21).
4.2 Second cycle

Data collection was mentioned more frequently in the second cycle’s State Report of 2007 (United Kingdom, 2007a). The UK used statistics and its data collection methodology (particularly referring to the 2001 census) as evidence of compliance with many of the FCNM requirements. In addition, under the articles relevant to education, the UK reported on its introduction of data collection in 2002 to show how it is monitoring attainment levels by ethnic group and matching this to other macro data. In its Opinion, the ACFC also increased its mentioning of data collection, including in the executive summary: ‘public authorities have taken steps to strengthen equal opportunities in their functions and recruitment practices, including through the collection of data on the situation of minority groups’ (Advisory Committee on the Framework Convention for the Protection of National Minorities, 2007: 1). The ACFC also praised the expansion of ethnic group categories for the 2011 census but expressed concern that this did not include Cornish (at that time not recognized by the UK as coming under the scope of the FCNM). Furthermore, the ACFC suggested that authorities in Northern Ireland, Scotland, and Wales needed to ‘pursue further their efforts to collect data on the situation of minority ethnic communities in all relevant sectors’ (ibid.: 18). This suggested that the data collection issues lay mostly with not collecting enough data, rather than having too much. Again, in its Government Comments, the UK’s response frequently refers to data collection – both as answers to flagged issues and to specifically address the aspects of Northern Ireland, Scotland, and Wales. Here, the UK referred to the forthcoming 2011 census and the more-accurate data it was expected to collect, as well as discussing the difficulties in the case of Scotland where minorities constitute a much smaller percentage of the population than elsewhere (United Kingdom, 2007b). The document also referred to many of the methods and population surveys outlined in Section 2 of this article. As with the first cycle, the general message is one of praise for data collection and a push for more, rather than any criticism or concern. Thus, the Resolution adopted in 2008 praised the UK’s data collection in general but stressed that improvement was needed outside of England – especially in Northern Ireland (Committee of Ministers, 2008).

4.3 Third cycle

Owing to the delay in submitting the second cycle State Report, the UK submitted its third cycle State Report only three years later, in 2010. Again, the report mentions data as a method of meeting the UK’s commitments under the FCNM and references are made to the (still)
forthcoming census in 2011. In addition, there are extensive sections where data collection is mentioned in relation to crime and education, as well as a greatly increased focus on Wales (United Kingdom, 2010). The ACFC Opinion adopted in 2011 mentions data collection in the second sentence of its executive summary, suggesting that it is one of the main points of relevance in the UK case. It reads: ‘Substantial progress has been made in promoting and advancing equality of persons belonging to different minority ethnic communities, based on extensive data collection’ (Advisory Committee on the Framework Convention for the Protection of National Minorities, 2011: 1). This is a direct suggestion that the situation of minorities in the UK is improving in large part due to the intensity of its data collection. It goes on to clarify that ‘even though data collection does still not manage to capture the situation of persons belonging to all minority ethnic groups, improvements in this field have led to better needs assessment’ (ibid.: 5). This acknowledges the progress as well as the limitations of data collection. Nonetheless, there are instances where the ACFC suggests that a lack of existing data is hindering policymaking and needs assessments, and thus includes in its recommendations that data collection should be increased. These instances include the context of Northern Ireland. Again, the ACFC points to the 2011 census and suggests that the extra improvements should enhance the existing knowledge of data disaggregated by ethnic group and the UK’s consequent ability to address ongoing issues. The Government Comments (United Kingdom, 2012) mostly followed the same path as the State Report and the Resolution adopted by the Committee of Ministers in 2012 stated: ‘The authorities have continued to collect an extensive range of data on the situation of persons belonging to minority ethnic communities in various areas of life, as a means to evaluate better the needs of these persons and improve policies and measures to combat both direct and indirect discrimination’ (Committee of Ministers, 2012). This represents a change from the second cycle’s Resolution which gave praise but also stressed the need for improvement; the omission of this latter aspect suggests an improvement had been made in this regard.

4.4 Fourth cycle

Finally, the most recent monitoring cycle saw the UK submit its fourth State Report in 2015, which again featured many references to data collection (United Kingdom, 2015). This was mostly given as evidence for how it was meeting its FCNM commitments. In some areas, it repeated aspects mentioned in past monitoring cycles such as issues relating to education, the health sector, and crime, plus a geographic focus on Wales. In addition, a focus was placed on the Gypsy, Roma, and Traveller communities in relation to the new category for the 2011
census, something the ACFC had previously highlighted. In its Opinion, the ACFC mentioned data collection more frequently than in any previous report. It featured many areas of praise but also drew attention to some aspects which it felt needed further attention; this subsection unpacks these in turn. The ACFC praised the 2011 census changes with regard to Gypsy, Roma, and Traveller communities and also the inclusion of the Cornish write-in box, but added that this latter element should have its own tick-box in the future (Advisory Committee on the Framework Convention for the Protection of National Minorities, 2017: 10).

For the first time, the ACFC has included an entire subsection on ‘Collection of Equality Data’ (ibid. p.17) under Article 4, which is concerned with equality and measures to promote this. In previous monitoring cycles, much of the data collection discussion had of course come under this Article; this cycle marks a significant increase in the prominence of the issue by including a separate concentrated subheading. Accordingly, this subsection discusses data collection in detail and summarizes the UK situation. Praise is again given for the broad collection of disaggregated data in the UK, but the ACFC expresses concern with Gypsy, Roma, and Traveller data as well as the general situation in Northern Ireland (ibid.). In its Comments, the UK government outlined their ongoing strategy for Gypsy/Traveller development and acknowledged the need to go further with their actions – similarly with Roma communities. Moreover, the UK government once again outlined how increased data collection would help meet these goals across varying sectors (such as education and housing), as well as recognizing that data collection was not as advanced in Scotland and Wales (United Kingdom, 2017). In addition to this, the Government Comments also announced an audit to ‘tackle racial disparities in public service outcomes’, which would review gaps in data collection as well as an intention to publish all data in one place (ibid.: 2). The outcome of this second point is the situation outlined in Section 3 of this article and the government portal for ethnicity facts and figures. Finally, the Resolution adopted by the Committee of Ministers (2018) reiterated the need for the UK to collect disaggregated data on Gypsies, Travellers, and Roma, and again mentioned the Northern Ireland issue. Interestingly, there was no specific praise for the UK’s existing approach as there had been in previous Resolutions.

4.5 Summary of the four cycles

Analysis of the four monitoring cycles shows that mentions of ethnic data collection have been consistently increasing in UK and ACFC documents since 1999. There was a significant rise after the first cycle, with the second and third cycles featuring much more widespread contexts
of data collection. The UK has consistently outlined that it uses data collection as a method for meeting its commitments under the FCNM, particularly in relation to Article 4. Throughout, the ACFC have given clearly worded praise to the UK for the depth and breadth of data collection disaggregated by ethnicity, and its recommendations usually only pushed for still more data. This focused mostly on demographic/geographic elements, such as extending data collection to other minority communities and increasing the scope in Scotland, Wales, and Northern Ireland to match the levels collected in England. The census of 2011 was a common theme throughout the second, third, and fourth cycles, with the ACFC recommending the inclusion of Gypsy, Roma, and Traveller communities as well as the Cornish. The former was fulfilled in the 2011 census with an additional tick-box, however the ACFC then called for this to be matched across all areas of data collection – not just the census. With the Cornish issue, the ACFC noted the significant number of individuals using the write-in option to specify Cornish identity and pushed for this to have its own tick-box in the 2021 census – particularly after the UK’s 2014 recognition of Cornish under the FCNM.

To summarize, the ACFC have mostly heaped praise on the UK for its data collection, clearly stating that this action has been successful in ‘promoting and advancing equality of persons belonging to different minority ethnic communities’. Furthermore, there is very little criticism at all regarding data collection and such criticism is directed mostly at recommending an increase in the process. Although the FCNM reminds states, in its thematic commentary, to ensure self-identification and avoid minorities self-collecting data, this is not mentioned in the UK context – presumably confirming that the ACFC is satisfied with the manner in which the UK conducts its data collection.

5. Conclusions

This article has demonstrated how widespread ethnic group data collection is in the UK. The historical overview showed how this developed over time throughout the 20th century, accelerating in the past two decades. One particular driver of this was the Equality Act 2010 and its PSED, which expanded previous legislation and provides a clear outline of the legal requirement of public sector authorities to collect data on ethnicity. As such, Section 3 of this article was dedicated to outlining some of the core areas in which ethnic group data were available, demonstrating the depth of this as well as the sources used to collect such data. This is so widespread that it becomes difficult to disentangle which areas of the UK applies to each outcome and where these data are sourced from. This stems largely from the blurry lines that
differentiate England, Wales, Scotland, and Northern Ireland, with the Gov.uk website not always stating which region is covered by the data. This aspect is picked up in the ACFC monitoring cycles, where the *Opinions* mention issues with Wales, Scotland, and Northern Ireland. In short, England has by far the most data. The analysis in Section 4 then examined the four completed monitoring cycles of the FCNM for the UK, looking at the perspectives of both the UK and the ACFC. This showed that the UK listed its data collection as evidence of its compliance with the FCNM across a wide range of factors, but with a particular focus on Article 4 of the FCNM. In its responses, the ACFC praised the level of data collection but also recommended that it should be further increased in terms of minorities covered (Gypsy, Roma, Traveller; Cornish) and geographical scope (Wales, Scotland, Northern Ireland). There was no criticism on the methodology or the areas in which the UK currently collects data.

Overall, both the UK government and the ACFC appear to strike a tone of ‘the more the better’ in their approach to ethnic data collection. Very little criticism is given to the methodology or number of areas covered; rather, it is justified as a means of meeting equality and anti-discrimination goals through supporting policymaking and further legislation. Although specific policies are not mentioned, it is clear that the UK has been using data to show how its existing policies have been meeting FCNM requirements in some cases. Although it cautions against over-reliance on data, the ACFC generally praises the UK’s approach and almost all its critiques concern a lack of (geographic or demographic) data in specific areas. Interestingly, the ACFC and Council of Ministers initially praise the UK (particularly in the second and third monitoring cycles), but by the fourth monitoring cycle, this is reduced/non-existent. This could suggest that the ACFC and Council of Ministers take it for granted that the UK collects such a broad array of ethnicity data.

However, it is perhaps somewhat concerning that the UK government and the ACFC neglect to mention the outcomes of this data collection in terms of concrete policy enacted. References remain indirect, stating that ethnic group data collection is needed to inform policy, without mentioning specific policies enacted as a result of new knowledge provided by data. The most promising area for this is in health where it appears that data are changing risk group categorization, but in education and crime, it seems that little has improved. Thus, there is a risk that concentrating too much on the collection and not enough on the analysis of ethnic group data distracts from tackling the issue. Additionally, as shown by the inaccuracy of NHS ethnicity data and the reversal of the school census policy, ethnicity data collection should be scrutinized further in terms of its purpose and what is necessary. The UK should not get
complacent and believe that more data equals better outcomes; neither should the ACFC only push for more data without consideration for what it is being used for. Thus, the monitoring cycles of the FCNM could be a tool for addressing this balance of collection and analysis.
Notes

1. This article focuses on the UK and its practices of collecting ethnicity data, as well as the purpose for which it is done. It is not within the scope to directly compare it with other states. Indeed, as the European Commission article on ‘Data collection in the field of ethnicity’ points out, ethnic data are collected to varying degrees across the EU with a number of different purposes for doing so, depending on the country’s context (Farkas, 2017). Please see the original European Commission article for more details.

2. Censuses have occurred every ten years since 1801; the exception is 1941, due to WWII.

3. ‘Abroad’ is defined as including ‘Islands in the British Seas’ and ‘born at sea’.

4. This refers to the England and Wales census as well as the Scotland census. It does not refer to the Northern Ireland census – something which is explained in further detail below.

5. ‘Any other’ is in each case followed by a write-in box. Interestingly, the 2001 census also included the term ‘British’ alongside other ethnicities, i.e. Black-British, Asian-British.

6. The order altered depending on the UK region, i.e. in England, English was first on the list; in Wales, Welsh came first, etc.

7. The situation of Northern Ireland and the Troubles is multi-faceted and cannot be covered in detail here. What is important is the consequences of sectarian/religious conflict, the peace process leading to the Good Friday Agreement in 1998, and what this means for collection of data on ethnicity.

8. The official name is the Belfast Agreement, but this text will use the more common term of Good Friday Agreement.

9. Indeed, the Office for National Statistics provides guidance for those wishing to collect data on ethnicity and suggests that the harmonized questions it has developed should be used in order to ‘achieve a harmonised approach in the data collection process that will allow for consistency and comparability of statistical outputs from different sources across the UK’ (Office for National Statistics, n.d. c).

10. The previous duties concerned race, disability, and gender; they were enacted in 2001, 2006, and 2007 respectively, but were all replaced by the PSED 2010 that accompanied the legislative change which formed the Equality Act 2010. As with above, Northern Ireland has separate duties (Equality and Human Rights Commission, n.d.).


12. An in-depth explanation of the methodology is provided on the Office for National Statistics website (n.d. d). See, specifically, Section 6 (‘How the output is created’) and Section 7 (‘Validation and quality assurance’).

13. The author obtained a physical copy of a Jobseeker’s Allowance form and a tax credit annual declaration form from England, however it was not possible to find these online as the digital system requires a log-in and personal details for a claim / declaration.

14. Each year, all state schools are required to conduct a census on their pupils to collect or update a wide variety of personal characteristics, which the Department for Education can then obtain. England, Wales, Scotland, and Northern Ireland all have separate administrations for this.

15. Academic achievement refers to different indicators depending on the age group. At the younger age, there are data on subjects such as maths, reading, writing, science etc. These data then become available at GCSE level (around age 16 years), A levels (17–18 years), and further education. For the full list, and to view this data, see: https://www.ethnicity-facts-figures.service.gov.uk/education-skills-and-training

16. An example of the survey questionnaire can be found online here: http://moderngov.southwark.gov.uk/documents/s43536/GP%20Patient%20Survey%20questionnaire%20example.pdf

17. The GP Patient Survey website is available here: https://www.gp-patient.co.uk/report?practicecode=J84008

18. The authors provide a detailed overview of published health research involving ethnic data collection in the UK through NHS databases and ONS surveys.

19. The Stephen Lawrence Inquiry followed the murder of black teenager Stephen Lawrence in London in 1993 and found that the Metropolitan Police Service had been incompetent and had issues of institutional racism. This led to a series of recommendations including that the police force should have better information and transparency. This fed into the context of the Equality Act in 2010 and underlined the need for better data collection as a means of monitoring public authorities’ competence and equal treatment. The full report can be found on the UK government website: https://www.gov.uk/government/publications/the-stephen-lawrence-inquiry

20. Previously, British Crime Survey (BCS)
References

https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000168008bb6e

https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000168008c62b

https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000168008c6c2

https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900016806a4811

https://rm.coe.int/16806fb9ab


United Kingdom. Comments of the Government of the United Kingdom on the Opinion of the Advisory Committee on the Implementation of The Framework Convention for the Protection of
https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=090000168008b7a8

https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=090000168008fa46

https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016805a8c52

https://rm.coe.int/1680703a30

https://schoolsweek.co.uk/dfe-ends-divisive-pupil-nationality-data-collection/

Whittaker, F. & Camden, B. ‘Pupils who are not white British told to send in birthplace data’, Schools Week, September 23, 2016. https://schoolsweek.co.uk/pupils-who-were-not-white-british-told-to-send-in-birthplace-data/

https://www.wwl.nhs.uk/Library/IGov/Leaflets_and_Posters/Corp%20004%20FT%20Ethnic%20Orig in%20Leaflet.pdf