



STUDENTS AS
CO-CREATORS

Students as Co-Creators Disciplinary Research Collaborations
Project Title: Secondary Data Collection and Initial Analysis
for Immigration Research

Prepared by Alexander Bolotov

School of Computer Science and Engineering

Academic Year 2022

1. Executive Summary

- Initial project team meeting on the project methodology.
- Discussion on the practicality of the project, project priorities. Creation of the “Project Journey” document.
- Splitting the project team into mini-groups to cover the areas of secondary data analysis in the areas of 1. Employment and labour 2. Education 3. Migration (Alex)
- Identification of target visits of the team members to explore local archives and data collections across the UK

2. Background and Aims

Context: It is known that a significant part of data intensive research are the activities to analyse existing data/datasets. As the team was interested in the datasets related to migration, we planned to understand what is available for us, how to develop an approach to searching/selecting/filtering data and what could be some initial analysis to judge if a specific dataset can be useful for some given problem and in what aspect. Moreover, is there any way to find implicitly contained information in the datasets that do not explicitly mention this target information.

Goals: To find out

- What are the best UK databases available for immigration academic research?
- Can these databases be integrated for use in the immigration research?

3. Out Approach. After some first steps of the analysis of the available datasets we found that it is incredibly difficult to gain any information on a specific topic, specifically if we want to target rather large time interval. For example, if we are interested in second/third generation of migrants employment issues, it is far from straightforward to determine where to obtain relevant data. Moreover, we also found that several datasets we have been analysing require very specific level of knowledge. Here, in particular, our “data” degree students were very helpful. Therefore, the team had a series of brainstorming sessions which resulted in the following outputs.

Problem-setting. There are many sources that provide information about immigration trends and immigrants socio-economic status in the UK. As part of our work on occupational attainment for immigrants, we came to realise that these sources offer different data in terms of types (numerical and textual), span and content. This makes examining key research questions on immigration issues and the economic outcomes of immigrants in the UK difficult as it depends to greater extent on the quality and availability of suitable data. **What can we learn from demographic data about the UK's changing demography of migration?** The UK's demographic profile has seen tremendous changes over the Twentieth and early Twenty First centuries. Most notably, the migration of populations from South Asia, the Caribbean and Eastern Europe highlight that broader currents impact on the composition of the UK's own population. What can we gain from the existing data sets on demographic composition and change over time in the UK by examining the ontological basis of the application of data categories and the usage of data in shaping public discourse on the UK's population.

Research Questions: Can we identify/connect how policy approaches (such as immigration legislation, labour laws, and equalities) are built on and feed back into the ontological basis of data on such areas such as Employment and Labour, Education, Migration.

Methodology: Therefore, the team has identified a broader aim of this research -- to identify datasets that best inform the problem setting above, looking at some dedicated areas such as Employment and Labour, Education, and Migration. This determined splitting of the team in three mini-groups to work on the datasets given below.

Identification of the DATASETS AREAS/SUBJECTS to look at and the parameters (variables) to consider.

Employment and labour

- which ethnic groups are working or are concentrated in different sectors of the economy in the UK
- business and trade data within the UK or between regions

- Income for ethnic/regional groups

Education

- Level of education according to ethnicity
- Type of schooling (private, comprehensive/state and grammar/state selective) according to ethnicity

Migration

- Macro data on inbound and outbound migration
- Place of birth in the UK
- Languages spoken
- Areas of concentration by ethnicity and migration
- Remittances between the UK, South Asia and the Caribbean

First steps – datasets analysis

The team found it useful to ask minigroups to fill the table below following the analysis of relevant datasets.

Target Data Topic	Description of data (What would it give us?)	Target Dataset (name & URL)	Open/Closed Access	Data Format	Main Features	Data Coverage (time, granularity, regions, countries,)
-------------------	--	-----------------------------	--------------------	-------------	---------------	--

The subsequent activities were devoted to

- First finding information to populate the table,
- Second – identifying prospective visits to the UK archives/data collections

4. Results

The following information was populated while answering the specific rubrics of the table.

Target Data Topic	Description of data (What would it give us?)	Target Dataset (name & URL)	Open/Closed Access	Data Format	Main Features	Data Coverage (time, granularity, regions, countries,)
This table presents information on Short-Term International Migration using three separate definitions of a Short-Term International Migrant	Nationality of short-term migrants by the primary purpose of their visit (only). Inflow and Outflow. Estimates from the International Passenger Survey, annual table.	Short-term international migration 01, Citizenship by main reason for migration - flows, England and Wales	Open (Office for National Statistics)	.xls Tables in different worksheets	Citizenship Year Reason for migration Direction of flow.	Migration towards UK. 2004 to 2019. Low granularity. 1 to 12 months migrants. 3 to 12 month migrants. Migrants satisfying the UN definition of short-term migrants
Nationality of migrants. Estimates of Long-Term International Migration, annual table.	Provides long-term migration data from different ethnicities. No reason for emigration/immigration provided	Long-term international migration 2.01a, citizenship, UK and England and Wales	Open (Office for National Statistics)	.xls	Citizenship Year Direction of flow.	Migration towards UK. 2004 to 2019. Low granularity. Long term migrants (1 year or more)
Regular job or profession of migrants entering or	Migrants occupation prior to migration	International Passenger Survey 3.12b, usual	Open (Office for National Statistics)	.xls	Year Direction of flow Occupation	Migration towards UK. 1976 to 2019. Low granularity

leaving UK by age. Estimates from the International Passenger Survey, annual table.		occupation prior to migration by age, UK			Age	
The primary purpose of migrants entering or leaving UK. Estimates of Long-Term International Migration, annual table	Long term main reasons for migration	Long-term international migration 2.04, main reason for migration, UK and England and Wales	Open (Office for National Statistics)	.xls	Reason for migration Inflow Outflow Year	Migration towards UK. 1991 to 2019. Low granularity
Data on non-UK nationals registering for a NINo	Adult overseas nationals entering the UK and allocated a National Insurance Number (NINo) on the National Insurance Recording & Pay As You Earn System (NPS). All adult overseas nationals allocated a NINo are included, regardless of their length of stay in the UK.	NINo Registrations To Adult Overseas Nationals Entering The UK	Requires registration	Downloadable in various formats	Nationality Gender Age Local authority for registration	2002-2021 Only UK coverage Low granularity
Provides an overview of the latest statistics on Sponsorship. More detailed data are available in the additional visas datasets.	Number of organisations on the register of sponsors for work Applicants for visas and extensions of stay for work using sponsorship certificates Applicants for visas and extensions of stay for work using sponsorship certificates	Sponsorship summary tables, year ending March 2021	Open (UK Government Website)	.ods	Working visa applications , Worker Type, Nationality	Work sponsorships 2010-2021 Low granularity
Work sponsorship (Certificate of Sponsorship) - Detailed Tables	Applicants for visas and extensions of stay for work using sponsorship certificates, by tier and industry type and nationality	Work sponsorship (Certificate of Sponsorship)	Open	.xlsx	Year Type of application Category of leave Industry Number of Applications Nationality	Work sponsorships 2010-2021 High granularity
Entry clearance visas	Applications for entry clearances visas and outcomes	Entry clearance visas	Open	.xlsx	Year Nationality Visa Type	Entry clearance visas 2010-2021 High Granularity

applications and outcomes. Entry clearance is the procedure used at British missions overseas to check, before a person arrives in the UK, if that person qualifies under the Immigration Rules for entry to the UK.	Includes work related Entry Clearance visas	applications and outcomes			Applicant Type Case outcome Number of applicants	
Local area migration indicators	Different migration-related data sources at local authority level including migration flows, non-UK-born and non-British populations, National Insurance number registrations, GP registrations, and births to non-UK-born mothers.	Local area migration indicators, UK	Open	.xlsx		Local area migration 2010-2021 Low granularity

The table above provided us with some useful classification of the datasets, main features of these datasets, and gave some understanding of the possible ways for the data interpretation and identification of new datasets (key words) in order to obtain implicit knowledge.

Visit to People's History Museum & John Rylands library archives (Manchester), 01 June 2022.

The team planned this visit to migration related places with the anticipation that it would be extremely useful for the project and would greatly contribute to our study of relevant materials/archives, artifacts/etc available at the places visited. We have chosen to visit two places in Manchester – People's Museum and John Rylands Library Archive. Manchester area is very representative in relation to various aspects of migration.

During the meeting with Lianne Smith, Archivist and Library Manager | Ahmed Iqbal Ullah RACE Centre Visit, we learned that John Rylands Library Archive is an extremely powerful place with great potential. At the same time, we learned that while South-Asian and Afro-Caribbean migration are well represented, the Eastern European Immigration Resources are limited. During the visit we had fruitful discussion on the means of usage of archives and discussed various issues related to the project, such as "scaling", in particular, if the migration ontology could be a route through the datasets. It became clear that two lines of scoping are opening – first around Archives and second – around Specific Datasets linked to communities. A particular task which could be of an interest to the community - bridging the gaps in local archives. The main lesson learnt is that the work on the resources, interpretation of data are crucial under the umbrella of "exhibitions come and go but resources stay", and migration ontology could be one of them.

The visit has been a very successful step towards building the relationships with the Archives, below is an email from the team leader at the Archive who we have met:

It was a pleasure to meet you last week, and I'm pleased to hear your visit was a fruitful and enjoyable one. I'm very much looking forward to continuing the conversations in due course.

5. Discussion

As we expected the project's area -- secondary data analysis -- is an important field but any type of the usage of secondary data requires a specific approach, understanding of the main parameters/variables which could lead to more efficient data discovery, and what is even more important, the interpretation of the data. A particular issue here is the identification of the ways to find implicitly contained information in the datasets that do not explicitly mention the targeted categories/problem aspects.

6. Conclusion and Recommendations

The project has brought some interesting and useful insights into the SDA, proposing some useful methodology for carrying on the analysis of the data which could be useful for future data intensive projects, specifically, for secondary data analysis related.

7. Dissemination

We plan to disseminate/discuss the project's results among the UoW Migration community.

8. Team Reflection

One of the main problems in our project management was related to changes in our student's partnership. Perhaps, as I often happens, some students who initially expressed their interest, and were active, found later that they overcommitted. However, those who stayed with the partnership and carried on work have produced fantastic contributions!