MANCHESTER 1824
The University of Manchester

Q-Step Data Fellows

2024

The logos indicate the types of places where our students spend their time as data fellows. They cover organisations in the public, private and voluntary sector.

q-step@manchester.ac.uk

www.humanities.manchester.ac.uk/q-step/

Q-Step Data Fellows 2024

The booklet tells the stories through a series of posters that each student is required to deliver at the end of their data fellowship.

Q-Step Data Fellows 2024

Who am I?

I'm Jackie Carter, Professor of Statistical Literacy in the Department of Social Statistics at the University of Manchester. I started the Q-Step Data Fellowships programme in 2013 and have led it since, winning accolades and awards for it. These include One in Twenty Women in Data, Women in STEM, Advance HE National Teaching Fellow and Principal Teaching Fellow.

These awards recognise the innovative nature of the Data Fellowship programme. The programme has also been recognised through the University of Manchester's Making a Difference annual awards.

What's Q-Step?

Q-Step was an initiative set up and funded by the UK government in 2013 to address the shortage of quantitative data skills that were being taught in social sciences and humanities university courses. Nineteen universities won funding to address this skills shortage.

The University of Manchester set up its Q-Step Centre in October 2013. We introduced more data skills modules into social science and humanities courses and started the data fellows programme to give students practice in using those skills in the workplace.

What is the Data Fellows programme?

Starting in the summer of 2014 with just 19 students, we have now placed over 350 undergraduates in around 60 organisations. The students are studying courses including criminology, sociology, politics and international relations, social anthropology, philosophy, economics, English language and linguistics and data analytics. Having learned some statistics in their modules and used real world data to empirically explore complex social problems, they can apply to be a Data Fellow.

If successfully selected - it's a rigorous process which includes an interview with the organisation they will be placed with - they are paid for up to eight weeks in the summer at the end of their second year to join an organisation to do some data driven work, often in the form of a research project.

Where do Data Fellows go?

All over!

The data fellows are hosted in organisation from the public, private and voluntary sectors. A student might spend time working in a national government department (like The Home Office of the Department for Education), or in local government (perhaps one of the ten boroughs in Greater Manchester).

Or they might be placed in a polling organisation or a data consultancy. Some go to university departments either to conduct research, or to assist with data analytics related to, for instance, differential attainment in degree grades.

Others spend time with media organisations (such as The Times/Sunday Times) exploring data journalism, and others may work for charity organisations (such as Greater Manchester Poverty for Action). Many stay in the north-west but we also have students placed in London and other major cities, and increasingly the data fellowship adopt hybrid working.

Our host organisations are prestigious, and we select them carefully to ensure the students have a good experience and are well-supported. Data fellows are often supported by former Data Fellows!

What do they do?

What they do varies widely and the best way to get a sense of this is to read the posters in this booklet. They are each asked to reflect on their learning across the summer and are given an Analytical, Research and Professional skills framework that I have developed to help them. They all learn skills that are highly in demand in the graduate workplace.

What do they do next?

When they return from the data fellowship we celebrate their learning – and the posters provide a way of them evidencing what they have learned. Many go on to use their new-found skills – and confidence in data analysis – in their third-year research dissertations.

Some stay in touch with their host organisations – some even continue working there. Many apply to do a Master's or PhD which enables them to further explore their subject, including often using quantitative methods. The further information section below gives pointers to some former data fellows who have entered policy and research roles.

Who pays?

Initially the data fellows were paid for by the Q-Step grant. When this came to an end the University of Manchester agreed to continue funding the data fellowship programme. We also have a model where if an organisation wants to take more than one student, they pay for the additional one(s) and we highly encourage this. We have more students than we have placements!

Further information

- *70% of the data fellows have been female. 25% have been from historically under-represented groups.*
- *The majority of data fellows have not studied mathematics beyond high school level.*
- *The book "Work placements, internships and applied social research" (Carter, Sage 2021) draws on ten case studies of which 8 are former data fellows*
- *These two open access booklets include stories of former data fellows.*
  - *Pathways into Research: Social Science Alumni Stories (2023)*
  - *Pathways into Policy: Social Science Alumni Stories (2022)*

CivED CIC, The Engine Room: Analysing youth flight in Southport

A DATA-DRIVEN APPROACH TO REGIONAL INEQUALITY AND ECONOMIC DIVERSIFICATION

Emily Bennett | BA Politics, Philosophy & Economics

Overview of the Data Fellowship

This project aimed to quantify and analyse youth flight from Southport, focusing on its impact on regional inequalities and economic diversification.

This required (1) cleaning and analysing data to compare young populations between Southport and neighbouring urban centres, (2) assessing the economic impact of these differences, and (3) developing policy recommendations to encourage multi-directional mobility patterns and support local economic growth.

Data Analysis

I used Lower Layer Super Output Area (LSOA) data from the 2021 census to create a dataset containing detailed demographic information for each area. I then used these factors to create profiles such as 'young professionals' who are individuals aged 18-34 who possess a degree and are employed in a 'professional' occupation.

Combined membership values for the concentration of young professionals across various Southport and Manchester, normalised between 0 and 1.

Each area was normalised on a scale of 0 to 1 across these factors.

A score of 1 indicates an area with the highest possible concentration of young professionals, while a score closer to 0 suggests a significant depletion. The final scores were averaged to produce an overall "combined membership value" for each area. The analysis was carried out for all areas in England and Wales, and between Southport and Manchester for a more detailed comparison between an area with high in-migration of young people, and one with a large depletion.

Findings

- *Coastal communities and deprived inner-city areas suffer from high out-migration of young people.*
- *Young people with A-level or degree-level education are more likely to move away from these areas than less-educated people.*
- *Young professionals are highly concentrated in urban neighbourhoods such as Ancoats in Manchester. These areas have good transport links, professional job opportunities and many amenities, and their housing stock is mainly apartments. By contrast, Southport is not well connected to any major cities, has a disproportionate reliance on leisure, retail and care workers, and has amenities which appeal to older people and families.*
- *The rise in popularity of remote working could attract more working professionals to relocate to coastal communities. Community working spaces and good internet connectivity would improve its appeal.*
- *Cultural regeneration is necessary to attract people to traditional coastal communities. Disused retail space could be repurposed for creative and leisure use. Successful regeneration projects have created vibrant social hubs consisting of mixed-use spaces that can be used differently in the daytime and evening and offer changing activities, exhibitions and entertainment.*

Key Skills Learnt

Through the course of the data fellowship I built on a range of research and professional skills, including:

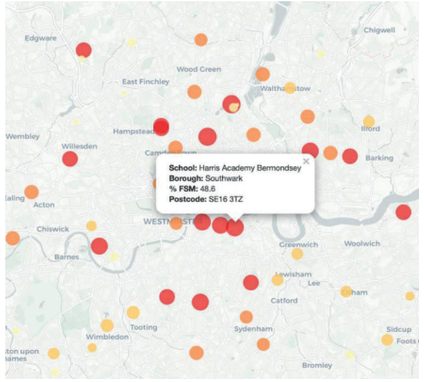
- *Deciding the best way of answering the research question, including knowing what data would be most relevant and which methodology would be the most useful.*
- *Cleaning and manipulating large datasets in R and Excel.*
- *Being able to draw conclusions from data analysis and knowing whether causality can be inferred.*
- *Managing my time effectively to balance deeper research with being on track to complete the project.*
- *Communicating effectively with the organisation and being able to ask for help and share ideas.*

The Engine Room: a co-working and co-learning space in Southport developed by CivED CIC aiming to produce a generative 'civic ecology'.

Girls in Data - Tracking the Impact of data related 'Inspiration Sessions' on Young Girls (20pts)

STRATEGIC INTERVENTIONS FOR GIRLS IN DATA FOR UK SCHOOLS

Jaya Dosanjh & Abigail Denner| BA Economics & BA Politics, Philosophy and Economics



Geospatial Mapping of All Girls Secondary Schools in London based on % FSM.

Overview of the Data Fellowship (23pts)

During our data fellowship with Girls in Data, we worked on diverse tasks supporting the charity's goals. You analysed government and think tank datasets, proposed an exhibit for the Florence Nightingale Museum on women's health inequalities, redesigned surveys, and helped set up a CRM system through Microsoft Non-profit 365.

We also evaluated the charity's educational sessions and developed a report on gender inequality in STEM subjects, using Free School Meals data to suggest expansion strategies for under-represented regions.

Data Analysis

This data fellowship allowed us to progress leaps and bounds with our data analysis. We had the opportunity to become confident working independently to select and analyse datasets using both R and Python coding to create accessible data visualisations including geomapping with the leaflets software.

We used these methods to conduct exploratory analysis to identify relationships that impact girls' ability and desire to remain in Data and Tech subjects.

Findings

- *Girls make up only 21% of Computing and 17% of Engineering students at GCSE, this indicates that girls drop out of Data and Tech subjects as young as 14.*
- *This statistic progresses to university, women make up only 18% of computing and 17% of engineering students.*
- *While girls report being 9.7% more confident than boys in Science, they are 5.7% less confident in Engineering and Maths.*
- *Women represent only 24% of the STEM workforce.*
- *From the top 10 local authorities in London by average percentage of free school meals (FSM) for all girls secondary schools, Islington leads with the highest % FSM, followed closely by Tower hamlets and southwark, indicating a high level of socio-economic deprivation in these areas.*

Key Skills Learnt

- **Communication & Collaboration:** *We both delegated tasks efficiently among each other and sought feedback from our mentors to improve project outcomes. As a result of this we effectively summarised tasks during group meetings and presented progress updates in our weekly catchups.*
- **Creativity:** *Our creativity was developed by enhancing tasks with engaging visuals, such as considering specific colours for our data visualisations, and redesigning feedback questionnaire forms to appeal to younger children using emojis and star ratings.*
- **Adaptability:** *During the course of this fellowship, we had to familiarise ourselves with new never before seen concepts and adjust to tasks based on this. There were many times this occurred such as applying The Common Data Model when creating Entity Relationship Diagrams, using new packages in Python and R-Studio such as leaflet and even learning the fundamentals of a Microsoft Non-profit CRM system.*
- **Networking:** *Our networking developed through the opportunity of meeting and video calling highly regarded professionals. This internship also presented us with the opportunity of attending Big Data London which expanded our industry knowledge further.*

Attending a talk at Big Data LDN.

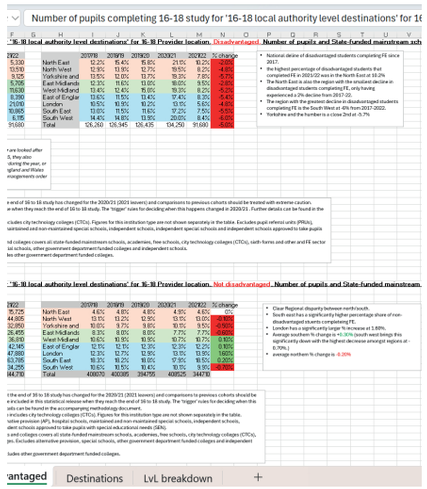
THE NORTHERN POWERHOUSE PARTNERSHIP

Eleanor Eyles | BA Politics and Modern History

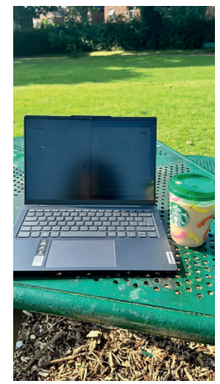
This Summer I participated in a Q-Step internship as a Data analyist for my host organisation, The Northern Powerhouse Partnership. NPP is a business-led think tank representing the voices of business and civic leaders across the north of England.

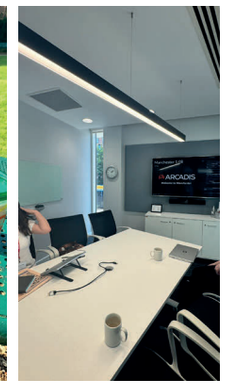
The photo above is of NPP's Net Zero Report launch as forgemasters in Sheffield. It was an great ooportuiny to meet policitcal leaders such as Oliver Coppard (Mayor of South Yorkshire) and Anne Handley.

The Net Zero report was written prior to my internship, but I was able to help work the event. NPP collaborates with the public and private sectors to advocate for increased investment and broader devolution throughout the north. Although originally set out as an economic project, NPP's ambition of transforming the northern economy is shared by companies, mayors, local politicians, and universities rather than being a central government economic policy.



During my 8 weeks at NPP, I worked on a research project centred around further education. Although we did not have a clear vision of what we were looking to discover, this project was primarily data-driven, in which we allowed the data to guide us. This included searching data playgrounds, The OECD data sets, and most frequently the government website.





The Data Set I looked into most was the enrollment rates across regions for further education. This included specifics like whether students were going into apprenticeships, vocational courses, or A-Levels and B-TECs. Other Variables included whether these students were of disadvantaged/ NON-disadvantaged status, Sex, ethnicity, etc.

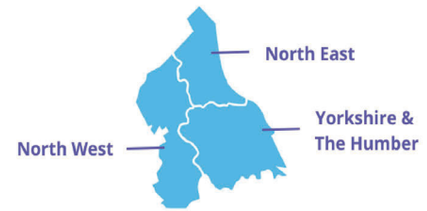
The Excel Sheet to the left shows the number of non-disadvantaged/ disadvantaged students completing FE across regions. This was one of the first pieces of work I completed during my tenure.

As NPP operates across the north, they have a primarily work-from-home structure, with monthly in-office meetings in any of NPP's offices. So although many of my days working looked like the image to the left, in which I would take my laptop and work from my local park, some of them also included visiting office buildings like Arcadis in Manchester and Addleshaw Goddard in Leeds.

Adapting from university to 9-5 life was tricky but I found the structure helped me manage my mental health in a positive way. As a history and politics student, raised in Bradford, this work is not only intellectually motivating but also personal.

Through my Q-Step experience, I've learned that although data research probably isn't the career for me, I've found a plethora of careers during my internship. After questioning my colleagues, I think I would thrive in a communications role, such as public affairs/ public relations in the political sphere.

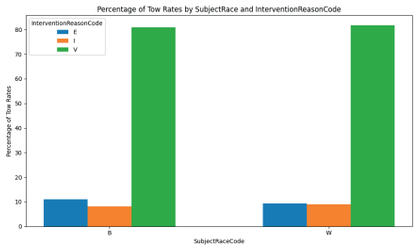
Overall, I had a lovely summer working on such a worthy cause and I feel very grateful to have worked with NPP on their further education report.



New York University – Marron Institute of Urban Management

ASSESSING RACIAL BIAS IN TRAFFIC STOPS

Amber Freeman | BA (Social Science) Criminology



A grouped bar chart for where I looked at the percentage of tow rates by race and the reason for being stopped in a bar chart.

Overview of the Data Fellowship

During my Data Fellowship I was hosted by The Marron Institute of Urban Management at New York University. This consisted of analysing 5 years worth of data (2018-2022) collected by police departments in Connecticut on Traffic Stops. My project aim was to assess to what extent there may be racial bias in traffic stops. In particular, looking at whether the law change on October 1st 2020 had an effect on racial disparities in traffic stops. The aim of the law change was to try and prevent traffic stops for no unwarranted reason, only being allowed to search a car if they have a probable reason to do so. To assess this question, I mainly looked at the tow rate of peoples cars and how this was affected by an individuals race, the reason for why they were stopped and the techniques used to stop that individual. I used Python programming language to code in Visual Studio Code.

Data Analysis

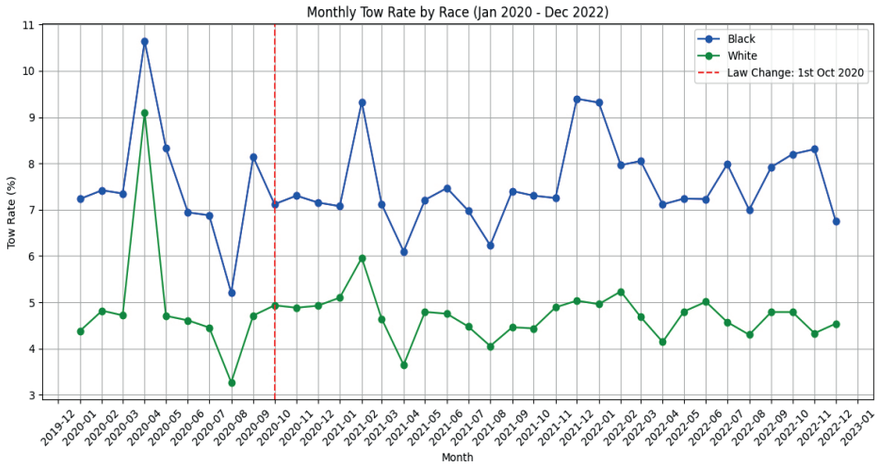
At the start, I looked at each datasets individually but eventually combined all 5 datasets into one single data frame to be able to identify patterns and trends across time. Through bivariate and multivariate analysis, I looked at basic descriptive statistics of the drivers race, ethnicity, age and sex, looking at how these changed when compared against other variables such as tow rate and the reason for being stopped. This involved me creating crosstabulations and converting counts into percentages. I filtered the data to show values for tow rate before and after the October 1st 2020 law change per race; presenting this in a line graph with a vertical line down the centre to represent the law change date.

Findings

- *I found that there was a higher tow rate percentage of people who are black, compared to white people. During some months, for example November 2022, almost double the amount of black individuals cars were towed (8.3%) in comparison to white people (4.3%).*
- *The graph I produced suggests it was more inconsistent for black people too, with larger, more frequent dips from January 2020 – December 2022.*
- *For both races, the most common reason for being stopped was for motor vehicle violations, contributing to 81.7% of reasons for being stopped for white people and 80.9% for black people.*

Key Skills Learnt

- *Manipulating, analysing and filtering data: this was a key skill which I had previously learnt on my University course, but during my data fellowship I developed it, using a different programming language; Python. I analysed and filtered data throughout the 8 weeks. An example of where I filtered data was when I wanted to look specifically at how an individuals race affected the reason for being stopped and if their car was towed away or not.*
- *Communication: whilst doing a data fellowship completely remotely and in different time zones, it made the reason for regular communication. Much more important. My skills of being able to communicate and ask questions confidently has grown massively over the 8 weeks.*
- *Time Management: I practiced good time management over the course of the data fellowship. The timetable consisted of having weekly meetings, where we would discuss the following week's tasks. This meant I had to efficiently manage my time so I could have tasks completed and therefore progress to the next ones.*

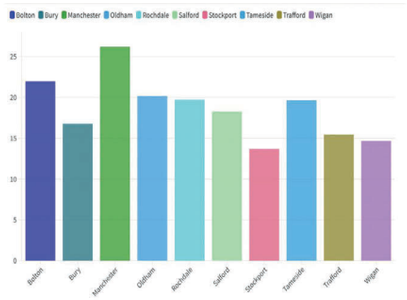


Line graph, including data from January 2020- December 2022, with law change line down the center.

North West Poverty Monitor – Resolve Poverty

NORTH WEST POVERTY MONITOR AND ADDITIONAL RESPONSIBILITIES AT RESOLVE POVERTY

Ellen Gallagher | Social Anthropology and Sociology



Percentage of people seeking debt advice in Greater Manchester local authorities in 2022 (featured in the Poverty Monitor)

Overview

At Resolve Poverty, the main project I have been working on is the Poverty Monitor, which is a free to access data resource which looks at poverty across the North West of England. It supports stakeholders to evidence poverty and it's causes using data shown in graphs, charts and maps. Thousands of researchers and organisations can use this Poverty Monitor to help tackle poverty.

The Poverty Monitor includes data across a range of different themes and categories, such as education, health, debt, labour market and social security. This data ranges from being at ward level, to local authority and national.

Data Analysis

I have interpreted and collected data-sets, and then determined and analysed whether they were useful or relevant enough for the Poverty Monitor. The data I collect is usually official data from the government, from sources such as StatXplore, ONS, or the official government website. I have to make sure the data is up to date, from an official source and is either at ward, local authority, regional or national level.

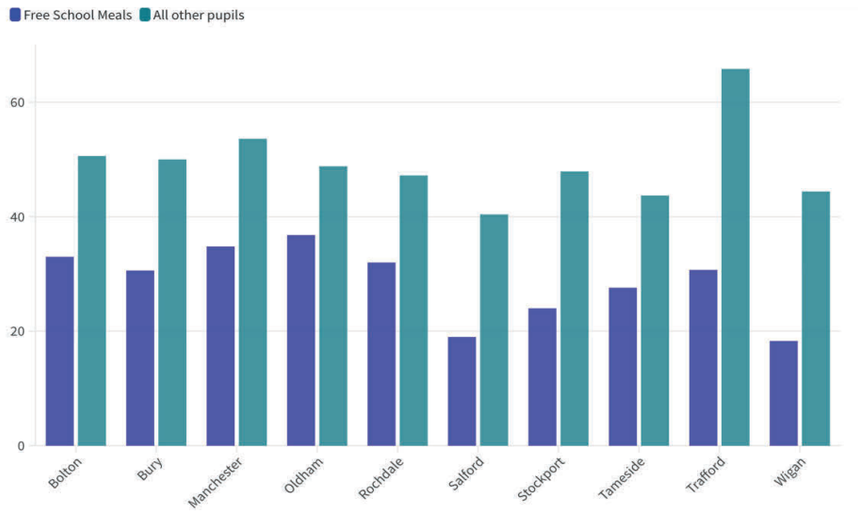
The data I search for is very specific, usually being under a category such as education and relating to poverty. It has to be up to date and focused on the North West region, preferably broken down to local authority level. The data I collect is then put into an excel spreadsheet and put into a graph ready for the poverty monitor.

Findings

- *The child poverty rate in the North West of England is higher than it is across the country as a whole.*
- *Two-thirds of working age adults and children experiencing poverty in the UK live in households where at least one person is in work.*
- *Report: Of the 29 county councils in England, 4 currently have an anti-poverty strategy in place. These are: Nottinghamshire County Council, Suffolk County Council, Surrey County Council and Warwickshire County Council.*

Key Skills Learnt

- *Learning about the ethics of collecting research.*
- *Time management and organisation.*
- *Advanced my knowledge in the following software: Excel, Zoom, Outlook.*
- *My confidence in communication and completing tasks improved.*
- *Learning how to be a self-starter and teach myself how to do tasks and learn software.*
- *Gained skills in searching for data and filtering it down to meet requirements.*
- *Research skills gained in searching for case studies, articles and data.*
- *Communicating data to an audience .*
- *Organising emails and logging responses into spreadsheets.*

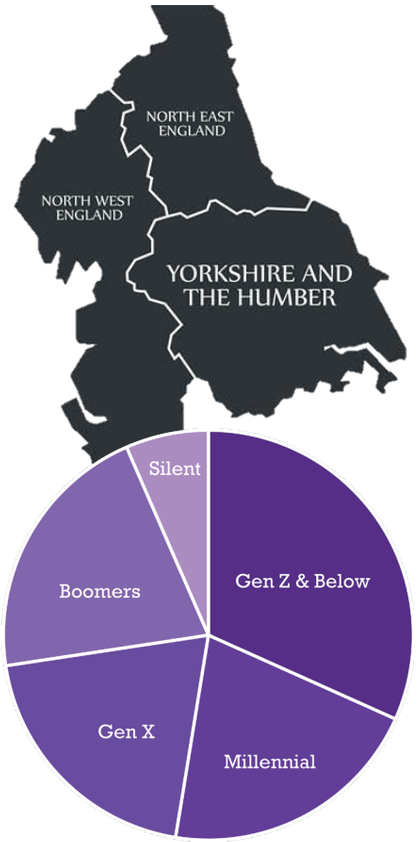


Percentage of FSM students who progressed to higher education in comparison to non-FSM students in Greater Manchester local authorities, 2021/22 (featured in Poverty Monitor)

IPPR North – State of the North 2025

MY Q-STEP EXPERIENCE – DATA ANALYSIS AND BEYOND

Daniel Galloway | Politics and Modern History



The North and it's Generations.

My Data Fellowship

My role was to begin the research for IPPR North's 'State of the North 2025'. The Institute for Public Policy Research North focuses on highlighting issues specific to the North of England, their annual 'State of the North' publication a key piece in presenting the condition of the North of England each year, and charting what needs to be done to improve the lives of its population.

Each year takes on a different theme, and this year was younger generations. Over eight weeks, I produced a series of briefings on the condition of young people in the North in relation to a range of issues including attitudes to democracy, health, housing, opportunity and culture.

Analytical Experience

My analysis largely concerned government data like the ONS and departmental publications, as well as research done by outside bodies like YouGov and Fields in Trusts. My work was intensive in locating relevant data for the issues on which each briefing was focused. Analytical skills were then required to understand how this data related specifically to young people in each of the Northern regions. The graphs, tables and charts I published then aimed to demonstrate this condition comparatively to the wider population.

It was also crucial that I framed the presentation of this data effectively. Findings were situated within the context of the wider literature, meaning data analysis had to be situated within wider problem-solving identification.

Skills: Data & More!

My data analysis skills have come on leaps and bounds this summer. I have developed my abilities in:

- *Data Location.*
- *Data Cleaning.*
- *Problem-solving using data.*
- *Analysis methods.*
- *Data presentation.*

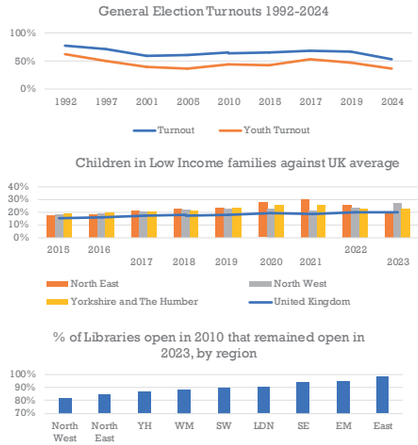
Beyond my work directly on data analysis, I had the great opportunity to attend IPPR events, and present my work to other individuals within the organisation, who have years of industry and academic experience. This developed my:

- *Networking Confidence.*
- *Argumentation and Persuasion Abilities.*
- *Presentation and Public Speaking Skills.*

My Q-Step experience exceeded my expectations and excited me about the prospects for the future.

Findings

While the report has many months of work to go until publication, many trends emerged within my research. The graphs below demonstrate some key findings from each of my research areas.



Funded by the Nuffield Foundation, ESRC and HEFCE

Understanding the impact of SHiFT

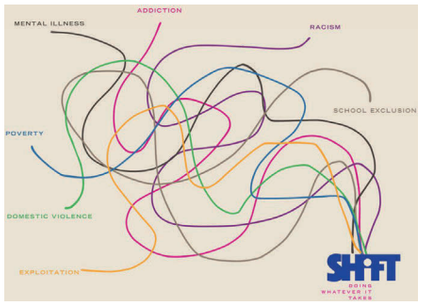
YOUNG PEOPLE AND CRIME: HOW TO SUPPORT THEM FOR SAFE AND BRIGHT FUTURES?

Xinran Han | BSocSc Sociology

Overview of the Data Fellowship

SHiFT exists to break the destructive cycle of children and young people caught up in, or at risk of, crime. SHiFT works in several practices in London, Greater Manchester, etc., which are made up of five practitioners (Guides) who each work with a small number of children for an 18-month period. The quantitative and qualitative data is collected and organised across four outcome domains: Loved and Cared for, Safe and Stable, Healthy and Happy, Power and Purpose.

My role as a data intern is to help SHiFT understand where we are most effective and where we have areas for development through analysing data from various professionals in youth justice and social care services and the Guide. I reached out to wider literature and data sets to contextualise the research and looked at different cross sections of the SHiFT population (ethnicity, age, deprivation, gender) to help SHiFT understand the outcomes.



Data Analysis

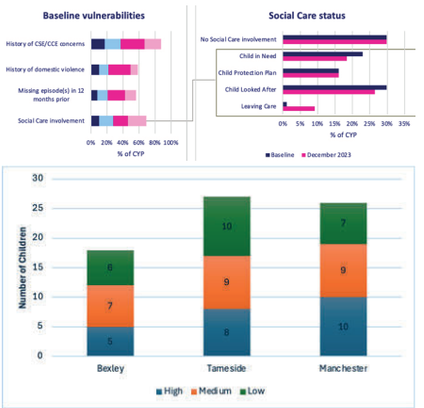
I prepared the data set for all the children we work with to analyse the outcomes. This involves checking the data with the online report done by the Guide and creating different spreadsheets for quantitative data and qualitative children's stories.

For example, to compare and analyse the baseline to the most recent 12 months of SHiFT for arrests, proven offences, missing episodes, and exclusions/suspensions, calculate the average, and make bar charts and line graphs by using Excel.

Findings

SHiFT has been curious about the effectiveness of in-person interactions between the Guide and the young person. I found relevant literature and hours comparison between social workers in general and SHiFT Guide, where the Guide has more hours spent with children. By analysing data of 71 children from 3 practices, I explored the contact hours, providing explanations for better intervention outcomes.

This report has shown that it is crucial to maintain accurate contact hour records, but a more detailed analysis of the way that Guide's time is spent is required, especially as the hours used here are self-reported. We also need to further understand the complex relationship between in-person hours and the quality of the relationship established between the young person and the Guide by analysing if outcomes for children are different for those receiving a low, medium, and high dosage and perhaps doing some additional qualitative work.

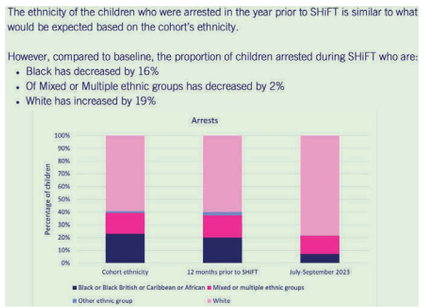


Key Skills Learnt

Throughout this fellowship, I developed my quantitative analysis and data visualisation skills, as well as the ability to search for previous research for literature reviews. I gained more confidence in using Excel, Canva, and PowerPoint. Furthermore, I improved my verbal communication skills as I worked closely with the data analyst and discussed my output with the national team. The national team praised this report as being very informative and helpful.

I also learnt about how SHiFT as a charity organisation produces an impact on helping children and collaborates with different services and professionals. Working with people with the same goal and hope makes my data fellowship very enjoyable, as I became a part of this positive change.

These experiences about charity work, especially UK child social care and youth justice systems, and therefore being able to figure out my passion is in charity to help to create a better world indeed.



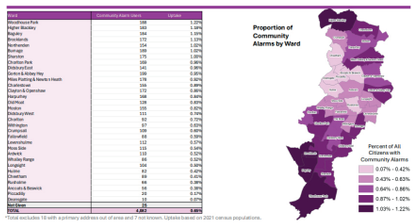
Example graphs for data analysis at SHiFT

Analysis of Technology Enabled Care (TEC) in Adult Social Care – Manchester City Council

UNDERSTANDING STRATEGIC DEVELOPMENT OPPORTUNITIES FOR TECHNOLOGY-ENABLED CARE

Chloe Little & Ian Wong | BAEcon Economics & Politics, BSocSc (Hons) Sociology

Community Alarm Uptake by Ward



Overview of the Data Fellowship

Tech-enabled Care (TEC) devices consists of the utilization of community alarm services and peripheral devices to uplifts independent living via preventive worsening of health. This method incorporates assistive technology to ensure TEC-serviced individuals can safely live in their own homes indefinitely, while also delaying the needed formal emergency care and support services.

The project looked at strategic opportunities to increase number of citizens receiving TEC by extending service to Manchester Wards imposing challenging demographics and geographical locations, in hopes of depleting inequality of residence TEC uptake. The completion of the 8-week data fellowship in the Council allowed us to perform variety of useful quantitative and qualitative skills. We had to conduct primary research and analysis of secondary demographics, and conveying our findings and recommendations in a coherent manner through a presentation to the head of social care and departmental senior stakeholders of MCC to contribute insightful data.

Data Analysis

As part of our data fellowship, we received advanced Excel training, and we were taught how to use ArcMap GIS. This equipped us with the skills to clean a dataset of 4,500 citizens. As the data we received was pulled from two systems we had to make a lot of decisions such as re-calculating citizens' wards from their postcode, and deciding which system provided more reliable information. After cleaning the data, we were left with 4,082 records suitable for analysis.

During analysis, we utilised complex nested formulae in Excel in order to extract various statistics on the demographic make-up of citizens receiving TEC. This involved integrating multiple datasets, such as those receiving adult social care in the city, and data from the 2021 census.

These insights allowed us to deliver visually appealing, visualisations of the data. Moreover, geospatial data was mapped across Manchester, achieved through multiple training sessions on ArcMap GIS. These visualisations demonstrated the depth of geographical inequality in access to TEC in Manchester, informing future strategic change in how this service is delivered.

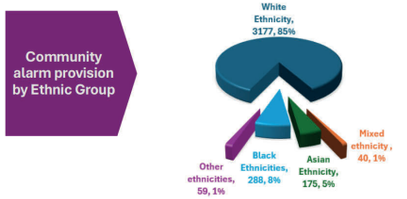
Findings

- *Of the 650,000+ citizens of Manchester, only a minor 0.65% of the population are enjoying the benefits of Tech-Enabled Care. The uptake percentage is much lower than Scotland (1.81%) in terms of Community alarm or Telecare.*
- *Community Alarm uptake concentrated in Wards furthest away from the Core of Manchester, with usage rate >9 times than Deansgate (20) and Piccadilly (10).*
- *69% of community alarm users inflicted with limb damage or mobility limitations, common in citizens age 60+ with poor physical health.*
- *Digital literacy strategy measures deemed highly inefficient: Tablets distributed to older citizens aged 70+ reached 4%.*
- *Pakistani and African groups are some of the largest ethnic groups in Manchester but respective population TEC uptake% remains at <0.2%. There are also no Roma people receiving TEC.*
- *Accessibility challenge of the TEC Smart Suite.*
  - *46.12% of health deprived households in Manchester do not own a private vehicle, increasing reliance on public transport for potential and existing service users.*
  - *Normal walking time (10 minutes) to public transport extended for Disabled (limb damage or Dementia afflicted) adults, Journey time reaching 1 hour of travel.*

Key Skills Learnt

As a result of working with real-world data, we learned a lot about the process of extracting, cleaning and organising data to make it useable, and about the various challenges that data analysts encounter during their work.

We gained valuable knowledge on how Manchester City Council operates, having the opportunity to meet with people across departments, explore the roles of different teams, and collaborate with and present to a wide range of employees. We gained significant presentational skills, delivering an hour-long presentation of our final findings to a meeting of almost 30 senior stakeholders across Adult Social Care in the council.

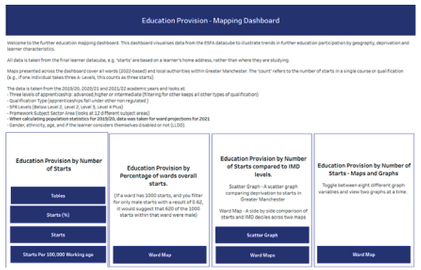


Pie Chart of Ethnic Group Distribution of Citizens with TEC

Greater Manchester Combined Authority – Analysing local trends in further education provision in Greater Manchester

VISUALISING AND MAPPING EDUCATION DATA WITHIN GREATER MANCHESTER

Atticus Morley | Politics, Philosophy and Economics



The image above is the homepage of the dashboard created, showing the source of the data and the navigation options available (in the blue boxes).

Overview of the Data Fellowship

Working with the Skills and Employment department in the GMCA to map further education trends within Greater Manchester. Over the summer I worked within the Greater Manchester Combined Authority (GMCA) office and used data from the Education and Skills Funding Agency (ESFA) to present a dashboard in Tableau to ease visualisation of data, helping further use and application of the data within the GMCA . Alongside this I did other data cleaning projects and a forecasting presentation around the potential of an increase in Manchester residents living under the national living wage . The projects undertaken all helped me understand how data can be applied in real world situations to benefit people around you.

Data Analysis

We Used data from the ESFA data cube between the years of 2019 and 2022. The data looked at the number of starts in education, so for example how many people in Rochdale started an apprenticeship in the year 2020. We split the data up between the; year, Levels of apprenticeship (advanced, higher, or intermediate), qualification type, subject area , gender, ethnicity ,age, and if the learner considers themselves disabled or not.

Through these filters on the dashboard an individual can discover their own analysis. A user can also filter between; total number of starts, starts per 1000 working age population, and % of all starts in that year. The dashboard offers bar charts, tables, scatter graphs, and ward maps. For example, when looking at the scatter graph area of the dashboard a user can select to compare all apprenticeship starts against the index of Multiple Deprivation (IMD) level in that ward.

Findings

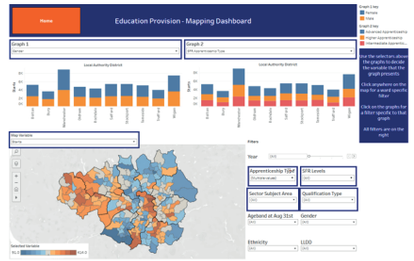
There are many different takeaways a user can get from the dashboard. The Scatter graph to the right shows the relationship between apprenticeship starts in Greater Manchester and the IMD levels of each ward. The lower an IMD score the more deprived an area is seen to be. As we can see there is a negative relationship between the two. Other findings from the dashboard have helped

visualise the divide between Manchester and the surrounding local authorities. Qualifications such as diplomas have been located within the Manchester Local Authority, whereas apprenticeships are heavily offered in the Bolton area compared to their other qualifications.

Key Skills Learnt

Throughout the Q-Step data fellowship, I gained valuable new skills. I was introduced to Tableau for the first time, and I expanded my proficiency in R. This experience has also strengthened my analytical capabilities and enhanced my ability to work with large datasets.

Most Importantly, completing the fellowship with the GMCA provided me with invaluable exposure to a professional, collaborative environment, whilst exposing me to strategic decision-making on a regional level and how cross-functional teams address complex social and economic challenges. This experience helped me build professional confidence, enhanced my communication skills, and taught me how to work effectively within a structured, goal-oriented organization.



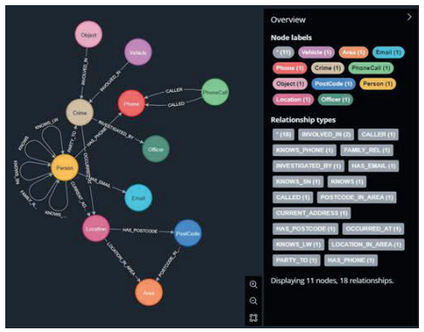
One of the Dashboards showing the filters and graphs that are available to users. The map is filtered for all apprenticeship starts.

The scatter graph dashboard showing a scatter graph with a line of best fit comparing starts of apprenticeships (X-axis) against the Deprivation level (IMD) (Y-Axis) . Lower IMD levels represent a more deprived area.

The Ditchley Foundation - Bridging Divides with Graph Networks

NEO4J AND UNITING PEOPLE INVOLVED WITH THE GREEN TRANSITION

Cassandra Papp | BA Linguistics



A schematic showing the types of nodes and relationships in a database

Overview of the Data Fellowship

At the Ditchley Foundation, I was first able to meet and work with a lot of other people interested in data analysis, many of whom originated from other universities or even areas of the world. Then, I was introduced to the three strands I would be working on during the summer: Quants, Existing Networks, and Networks Expansion.

Both of the networks strands allowed me to gain familiarity with various aspects of people the foundation was already in contact with and help the organization in locating additional people that Ditchley should contact in the future. The quants strand dealt with analysis of the types of people already present in the Ditchley Foundation's neo4j network.

Data Analysis

While working with the quants strand, I was tasked with looking into the geographical information of the people that were currently present in the foundation's neo4j database.

I first, however, had to start by working with some sample data (shown above) to fully understand how to use the coding language. In addition, working on the networks expansion strand gave me the chance to construct my own neo4j database.

This was based on twitter data of a collection of important people involved with the green transition here in the UK. I was then tasked with taking all the twitter users they were following to create a product that might help the Ditchley Foundation know who to interact with in the future.

Findings

- *In working with the quants data, my team and I discovered that there was a large bias towards people associated with the United Kingdom. We then showcased some of the most undervalued locations in the database for the foundation to improve outreach to.*
- *In searching through the data from my newly constructed neo4j twitter database, I was able to eliminate some of the more highly followed people (such as President Barack Obama) in order to identify who in the dataset seemed to be underrated. I was then able to create a custom dashboard which allowed the foundation to filter according to their own needs.*

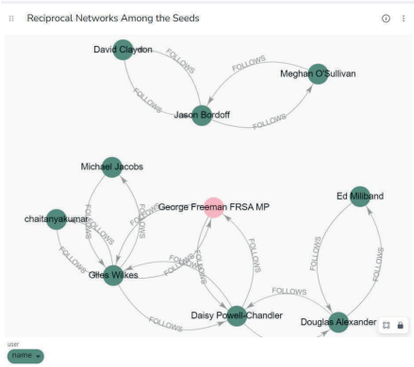
Key Skills Learnt

One of the most crucial skills I developed during my time with the Ditchley Foundation was how to showcase findings of data analysis. While I may have known what exactly a query of the database was showing, that information would remain useless to the less code-savvy members of the team unless I could showcase it in a way that was intuitive and easy to read.

Throughout the data fellowship, I was able to reformat tables and charts to make the outputs and the effects easy to follow without the effects shown being exaggerated in any capacity. Furthermore, I was able to introduce elements of interactivity that allowed those unfamiliar with how to tweak the code to still be able to have a say in the outputs produced.

In addition, I learned how to work well in a team with others and feel confident leading in my aspects of a project. Throughout my time at the Ditchley Foundation, I worked in three teams, one consisting of only two people (including myself), one consisting of four people, and one consisting of six people.

Learning how to stand up for my ideas as well as listen to other people's ideas looked very different in each of these settings, and I found navigating this to be very beneficial to my professional development. I know I will be able to apply these skills in future professional settings.

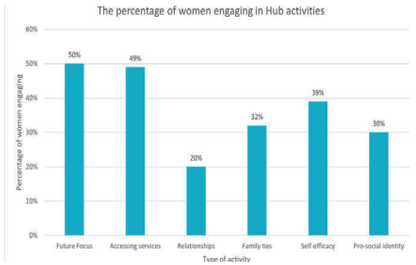


Reciprocal relationships between people associated with the green transition

Understanding the link between programme activity and an offender rehabilitative journey – HM Prison and Probation Services.

THE LINK BETWEEN ENGAGEMENT WITH CFO ACTIVITY HUBS AND OUTCOMES FOR FEMALE PARTICIPANTS.

Ella Parr | BA Criminology



Overview of the Data Fellowship

During the 8 weeks of my Data Fellowship, I worked on analysing data collected by HMPPS CFO (Creating Future Opportunities). My project explored the impact of the CFO Activity Hubs on female offenders across the UK. The aim of the Hubs is to provide participants with a safe environment that acts as a community, where they can engage with activities and improve their skills to have crime-free futures. I focused on three Hubs located in Sheffield, Bristol and Medway, and assessed what the different Hubs were doing to produce positive or negative outcomes for women.

Data Analysis

I used a dataset gathered by the CFO team which included a survey with 13 questions where the participant's Support Workers answer on a scale of 1-5, the likelihood the participant is to improve in this area. The dataset also included binary variables displaying which activities each participant engaged in, and variables showing the characteristics of each individual, such as ethnicity and age.

- *I mainly used Excel to analyse the data, utilising functions such as t-tests and ANOVA tests to determine statistical significances, and pivot tables to uncover trends in the data.*
- *I produced visualisations of my findings using excel, specifically tables and graphs.*
- *I also used other tools like Canva to develop infographics which present the key findings of my analysis clearly.*

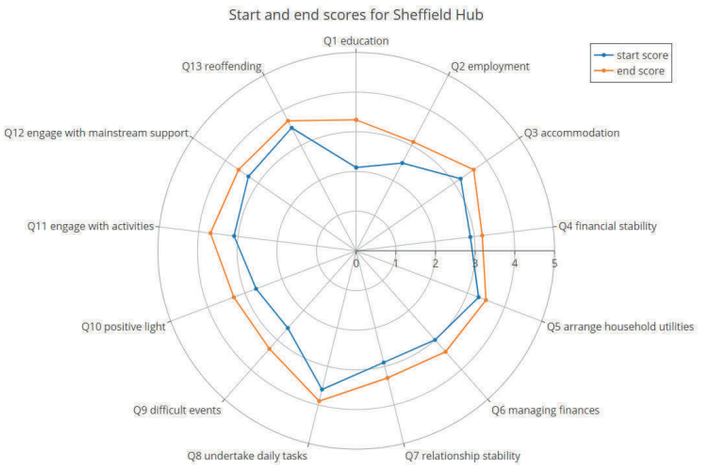
Findings

Firstly, I found that women have a less prolific offending history than men, they were more likely to have dependent children and be lone parents, and they faced more challenges with their personal attitudes. When looking into the types of activities women engaged in, I found that certain activities had more positive outcomes such as 'relationships' and 'self-efficacy'. Whereas other activities like 'family ties' had no clear benefits for women.

Finally, when exploring the outcomes at three particular Hubs, I established that the Sheffield Hub was extremely successful in comparison to Bristol and Medway. The Bristol Hub faced challenges as 52% of the women had served a custodial sentence, which results in issues with accommodation, relationships, recidivism and positive attitudes.

Key Skills Learnt

- *Communication skills – Presenting my research during an oral presentation to the CFO team allowed me to develop my public speaking skills.*
- *Analytical skills – I built further on my Excel skills and developed my critical thinking skills which allowed me to interpret correlations to determine my conclusions.*
- *Professional skills – working in an office full time was beneficial as I learnt professional skills like time management and teamwork. My project was largely independent which allowed me to work on my problem-solving skills when I came across any issues.*
- *I gained an insight into the work CFO do at the Activity Hubs as well as at HMP Risley on their Discovery Wing. I got to visit the prison and the Stoke Activity Hub, where I spoke to some ex-offenders. People shared their stories and told me how the Hub had changed their lives completely, which was so motivating to hear as my research will hopefully be used to develop the success of the Hubs in the future.*



Manchester City Council – Understanding prosperity: Developing a city-wide index

DEVELOPING MANCHESTER'S PROSPERITY INDEX: A NEW MEASURE FOR ECONOMIC AND SOCIAL WELL-BEING

Tharani Ponnamperuma | BA (Hons) Economics and Data Analytics

Overview of the Data Fellowship

During my data fellowship at Manchester City Council, I worked on developing a Prosperity Index aimed at supporting Manchester's economic strategy.

This index intends to provide a comprehensive view of economic and social well-being across the city's wards, helping policymakers track local progress and address socio-economic disparities.

My role involved defining what prosperity means for Manchester, identifying key metrics, and developing a methodology for measuring and analysing prosperity at a granular level.

Data Analysis

I analysed a combination of quantitative and qualitative data, including economic indicators (income, employment), social metrics (health outcomes, crime rates), and environmental factors (air quality, green space access). Using GIS and Excel, I created a database of indicators, mapping them to Manchester's wards and ensuring the index was accurate, current, and relevant to the city's unique socio-economic landscape.

Findings

- *Socio-Economic Disparities: The index revealed significant variations in prosperity across Manchester's wards, with disparities in income levels, employment rates, and access to essential services. Wards with lower prosperity scores often correlated with higher crime rates and lower educational attainment.*
- *Environmental Indicators: Air quality and green space access emerged as critical factors in determining overall well-being. Areas with greater environmental resources tended to score higher on the index, highlighting the importance of sustainable urban planning.*
- *Health and Social Cohesion: Health outcomes and social factors, such as community engagement and crime rates, were strong predictors of ward-level prosperity. Wards with stronger community networks generally displayed higher levels of prosperity, underscoring the role of social cohesion in economic well-being.*
- *Guidance for Policy: The index serves as a tool for targeted interventions, providing city officials with a clearer picture of which wards require specific resources and support to bridge socio-economic gaps.*

Figure 1: Snapshot of the prosperity index metrics selection

These findings underline the interconnected nature of economic, social, and environmental factors in shaping a community's prosperity and offer actionable insights for Manchester's economic strategy.

Key Skills Learnt

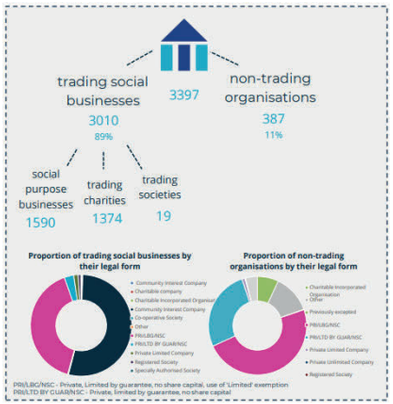
This fellowship enhanced my analytical and professional skills, especially in real-world data application for policy-making. I developed a nuanced understanding of socio-economic metrics, improved my data visualization capabilities, and gained valuable insights into the workings of local government.

Engaging in council meetings and workshops, I saw first hand how data influences decision-making, shaping my career ambitions towards economic strategy and policy analysis.

Economic Mapping: Bridging data and policy for regional and local transformation – Centre for Local Economic Strategies

HEALTH & LIFE SCIENCES SOCIAL ECONOMY SECTOR ANALYSIS

Simrin M Rahman | BA (Hons) Politics, Philosophy and Economics



Overview of the Data Fellowship

The Data Fellowship involved a comprehensive analysis for the Boosting Life Sciences Social Economy (BLSSE) project, led by several London boroughs and MedCity, with funding from the UK Shared Prosperity Fund. Tasked to support inclusive growth in health and life sciences, the fellowship required meticulous data analysis and cleaning to evaluate social economy organisations in this sector. This analysis encompassed both supply-side data, detailing social economy organisations, and demand-side data, highlighting local authority spending. The project aimed to provide a clearer picture of local support for social enterprises and enhance the capacity of these organisations in the health sector.

Data Analysis

The data analysis involved systematically collating, cleaning, and evaluating information on social economy organisations in health and life sciences across London boroughs. Supply-side analysis integrated data from Companies House, the Charity Register, and DataCity, with R scripts used to merge and refine databases.

Real-Time Industrial Classifications (RTICs) offered detailed categorisation, especially in fields like biotechnology. Demand-side analysis assessed borough and NHS Trust spending to quantify support for these organisations. Challenges, including differing trading and legal names, were managed by collaborating with data providers, ensuring dataset accuracy. This method enabled comprehensive cross-referencing by sector, legal form, and turnover for reliable conclusions.

Findings

- *The analysis identified nearly 3,400 social economy organisations across eight London boroughs, with 89% comprising trading entities such as community interest companies and trading charities, while 11% were grant-dependent, non-trading foundations or trusts.*
- *Camden and Westminster had the highest concentration of these organisations, while Hammersmith and Fulham showed comparatively fewer, with many organisations located in economically deprived areas, highlighting their role in community support.*
- *Findings showed a strong presence of organisations in non-residential care and mental health support, with high densities in deprived areas and varied organisational types across boroughs. Graphs highlighted a predominance of small enterprises, providing valuable insights for local authorities to boost economic inclusivity.*

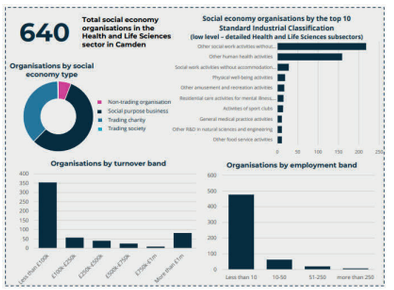
Key Skills Learnt

During the fellowship, I gained essential technical and professional skills that greatly enhanced my capacity for data analysis and collaboration. My proficiency in R developed significantly, particularly in data cleaning—a fundamental yet challenging aspect of analysis.

Learning R through trial and error, I became adept at troubleshooting with minimal R-specific supervision, which developed my critical thinking and resilience. For instance, resolving discrepancies between Companies House and DataCity data taught me the importance of thorough validation and attention to detail when managing large datasets.

Working at CLES also refined my communication skills. Regular updates on weekly stand-up channels helped me clearly articulate my progress, while engaging with colleagues' insights across different projects enhanced my teamwork. Networking furthered my understanding of community wealth building and local economic development, as I engaged in one-on-one discussions with team members.

Additionally, exposure to projects like the Korean government water company consultation allowed me to apply cultural knowledge and language skills, broadening my perspective on international economic issues. Observing the impact of social enterprises through engagements like the Coalfields Regeneration Trust highlighted their role in local communities, collectively strengthening my adaptability, problem-solving, and communication skills in data-driven economic work.

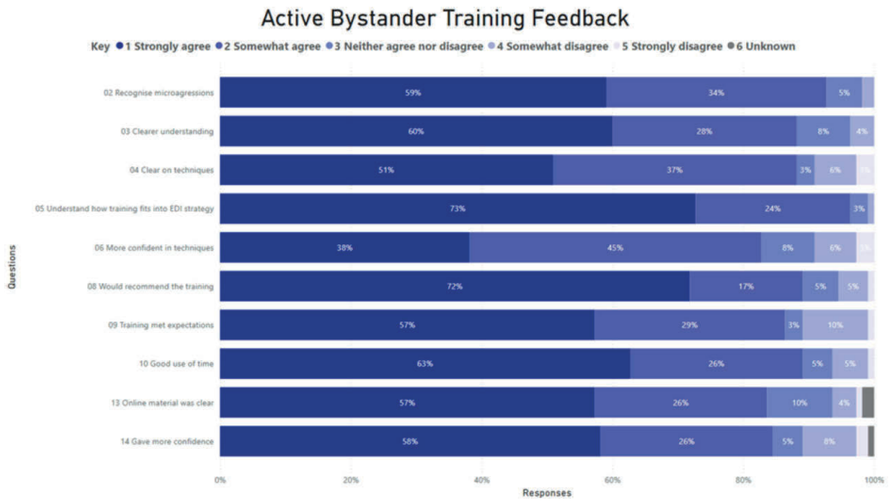


One section of the borough-by-borough analysis – a snapshot of the social economy in Camden

Equality, Diversity & Inclusion Directorate, University of Manchester

DATA FOR DIVERSITY: INSIGHTS & IMPACT

Illiyyin Aufa Syahrizal | BSocSc Politics & International Relations



A horizontal stacked bar chart summarizing the responses and attitudes to the success of the Active Bystander Training launched by the EDI Directorate.

Overview of the Data Fellowship

During my fellowship, I undertook two main projects and a series of analytical tasks. These projects involved presenting live data analysis sessions for university staff and analysing survey feedback from an Equity, Diversity, and Inclusion (EDI) training program. I applied data analysis skills to real-world issues, enhancing my technical abilities, communication, and adaptability through hands-on experience.

Data Analysis

Project 1: Presentations on Staff Diversity

For my first project, I analysed the diversity of protected characteristics within university staff along with and in comparison to the most recent available data on the global population, in alignment with the Equality Act. I cleaned and prepared the data, ensuring sensitivity and confidentiality. I used Excel for data handling, mastering pivot tables and various data protection techniques, and created clear, digestible visualizations. I also learned to use R, where I produced an interactive, colour-coded map showing staff ethnicity distribution, providing a compelling visual to represent the data's insights.

Project 2: Active Bystander Training Feedback Analysis

In my second project, I analysed both quantitative and qualitative feedback from the university's Active Bystander Training program. This required rigorous data cleaning and transformation. I utilized AI tools and personal thematic analysis to categorize responses, which highlighted participants' experiences. My analysis and reporting helped summarize training outcomes and provided insight into future improvements.

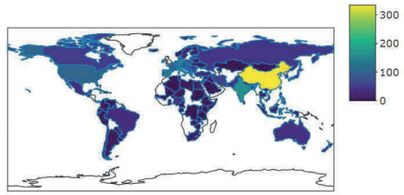
Findings

Protected Characteristics: Identified demographic trends among staff, with results presented to university staff, making the data accessible to those unfamiliar with data analysis.

Survey Feedback on EDI Training: Found that the Active Bystander Training had a positive impact on participants' attitudes and behaviour. My presentation of both quantitative and qualitative data underscored the program's effectiveness and areas for improvement.

Key Skills Learnt

- **Data Sensitivity & Protection:** *Gained experience in handling sensitive information by applying confidentiality measures in Excel.*
- **Data Visualization:** *Created clear, audience-friendly charts and interactive visuals using Excel, R, and Power BI. Learned the importance of colour-blind-friendly design for accessibility.*
- **Coding:** *Basic proficiency in R and Python, enhancing my data visualization capabilities and giving me confidence in using coding languages for complex analysis.*
- **Soft Skills:** *Improved my confidence in public speaking, managing audience interactions, and answering challenging questions on data quality. Collaboration and communication flourished, particularly in adapting based on feedback.*



My favourite achievement: learning how to create an interactive, colour-coded map using R which would provide the staff data for ethnicity when the cursor hovers over a country.

HM Prison and Probation Service Creating Future Opportunities

ASSESSING THE LINK BETWEEN PRISONER WELLBEING AND A CALMER PRISON ENVIRONMENT IN THE DISCOVERY WING.

Iris Potter Williams | BSc Sociology

Overview of the Data Fellowship

My data fellowship took place in HMPPS Creating future Opportunities. I had the opportunity to explore questionnaire data from phase 1 of the 'Discovery wing' at HMP Risley that prepares category C prisoners, for release through resettlement activities like managing finances or employability skills. Focusing on improving key performance indicators like wellbeing, resilience or practical problems like substance abuse or employability. Overall, aimed at creating calmer prison environment with the end goal to prevent re-offending, and improving employment.

Data Analysis and skills I used

The data was taken from self-complete questionnaires that test key performance indicators from start to end after completing phase 1. The dataset comprises of ordinal variables on a metric scale 1-5 such as resilience, impulsivity and wellbeing.

Data Skills I used

In excel to prove statistical significance, I utilised a range of tests of probability e.g. ANOVA, T-tailed test and chi-squared tests. Also using pivot tables to test certain trends in data.

Formulating a range of infographics and graphs through excel and Canva to present my work e.g. bar charts, radar (Spider) charts and tables that represent my data in lots of interesting ways.

Findings: Data is mainly formulated from veterans and/or sex offenders. My assumptions originally, was that due to literature, these participants would more mental health issues detected than counterparts. However, the difference was only 4-6 percent points and was not statistically significant. This implies mental health issues is a general offender trait, and rather than limited to one characteristic. Focusing on this 76% is hugely important.

Part 1 explored the start mean scores for people with and without mental health issues detected, finding people with mental health issues were falling behind in nearly all categories. **E.g. Hope averages 3.4/5.00. Whereas people without average 3.9. similar with wellbeing, resilience and impulsivity**.

Presenting the need to focus on progress for these participants and to implement special considerations to improve those with mental health detections as needing more support. **Part 2 looks at the progress made from start to end for people with mental health issues detected.**

My findings demonstrate significant improvement in resilience by +0.3 and a decline in impulsivity by -0.5, both statistically significant. Resilience suggests modules in confidence building', 'goal setting' and week 7 focus on 'resilience and mental health' can be linked to the positive impacts. Declines in impulsivity can be linked to modules like 'money management', 'goal setting' and 'managing on the outside world' are likely positive improvements. Decreasing impulsivity has implications on improving stability and problem solving and further can be linked to decline in substance abuse.

Part 3 focused on delving into distribution analysis and the effects of the programme on practical problems. Both 'resilience' and 'impulsivity' the impact of phase 1 appears to be successful. Both are associated the improvements in substance misuse, behavioural issues and associations with self-harm in ACCT referrals. Also, in improving holding down a job, gaining employment and securing suitable housing.

The programme brought significant improvement to participants overall.

Distribution analysis identified themes: participants with the most benefit from the programme, had self-reported issues, were provided with more tailored care, improved in wellbeing and hope. External factors like drug referrals needed more future analysis to understand the overlaps with the programme. Allowing me to make key recommendations around improving attrition rates or encouraging more relationships support like holding more family days.

Distribution Analysis: Hope	Start	End	Diff
Lived alone prior to coming into prison	3.4	3.7	0.3
Relationship problems	3.2	3.5	0.4
Wanted help holding down a job	3.1	3.6	0.5

Key skills learnt

- **Development of analytical skills and data manipulation:** *using excel, and ranges of probability tests like ANOVA tests, or t-test to identify areas of improvement and declines in the data.*
- **Critical distribution analysis**, *determining recommendations of the gaps in the programme that are transferrable to HMP Risley and other wings around the UK.*
- **Expanding my professional skills** *like collaboration with both statistical research team, and my intern partner Ella. Adaptability to changes in workplace, and data structures. Time management and working to structured schedules and deadlines.*
- **Ranges of data visualisation techniques and infographics e.g. radar charts.**

#03 Institute for Public Policy Research - Green Industrial Strategy

WORKING AS A DATE FELLOW AT THINK TANK IPPR

Ziyi Yang (Lily) | Sociology and Data Analytics

Knepp Rewilding Center Visit 2024

Overview of the Data Fellowship

During my data fellowship project, I worked as a researcher on two main projects: a social leasing project and a child poverty project. In the social leasing project, I conducted a literature review and gathered data from sources like existing studies and the Office for National Statistics (ONS). My goal was to compile this data into an organised Excel sheet for easier access and use by colleagues. The project aimed to explore ways to introduce affordable electric car leasing schemes for low-income households, promoting greener transportation while offering cost-effective commuting options. This experience helped me sharpen my skills in literature analysis and data organisation. In the child poverty project, I helped develop an interview guide for a study focusing on child poverty and migrant communities.

The guide had to be carefully crafted, ensuring it conveyed confidentiality and was easy for participants to understand, while avoiding language that could potentially be sensitive or triggering. This task improved my ability to communicate research concepts clearly and empathetically.

Overall, I gained hands-on experience in research methods, data management, and writing effective research tools, all of which have enhanced my ability to work on impactful social projects.

Data Analysis

Throughout this research, I primarily used Excel to analyse child poverty data and categorise cities by local authority. This revealed specific areas that require updated public policies and heightened attention to address child poverty effectively.

Knepp Rewilding Center Visit 2024

While I had previously learned Excel at university, this project reinforced the importance of being vigilant about the sources of secondary data, especially in comparative research. I learned that understanding the definitions, research aims, and context behind data is crucial—even when datasets seem to cover the same topic.

Key Skills Learnt

Communication: Communication has been one of the most valuable skills I developed during this data fellowship. My manager was often busy with tight research deadlines, which meant I occasionally found myself without immediate tasks. Instead of waiting idly, I took the initiative to reach out to other researchers at IPPR (with my manager's approval). This proactive approach led me to contribute to the research on Child Poverty and Migration, demonstrating my ability to seek opportunities and effectively communicate my willingness to collaborate.

Collaboration and Teamwork

Throughout the Child Poverty project, I ensured my work was transparent and well-coordinated by sending daily summary emails to the team. This helped us stay aligned on progress, ask questions, and improve our collective research. I also fostered an open dialogue by sharing any concerns or questions with the team, which enhanced both collaboration and the quality of our findings.

Manchester City Council-Knowing our residents: Population Segmentation

DEVELOPING POPULATION SEGMENTATION MODEL: AN EFFECTIVE WAY TO KNOWING RESIDENTS

Xingyou Yi | BSocSc (Hons) Sociology

Overview of the Data Fellowship

I focused on developing the population segmentation model in my 8-week data fellowship with Manchester City Council's Performance, Research, and Intelligence team. This population segmentation model aimed to categorise residents based on various indicators, such as well-being, to tailor policies that address each group's needs.

My tasks included data analysis, researching similar approaches by other city councils to identify gaps and opportunities for the team, and creating presentations for stakeholders to demonstrate the effectiveness of the population segmentation approach.

Data Analysis and Findings

For data analysis, I primarily used Excel and ArcMap GIS. My focus was two-fold: first, using ArcMap GIS to present demographic differences across wards, and second, using Excel to verify the validity and coherence of different datasets. Through both data analyses, I have found the disparities between different groups, such as universal credit, death age, income, and knowing the procedures of making targeted policies, which can be helpful to have a positive impact on different groups.

- *Among Manchester's 32 wards, characteristics of residents vary, leading to differing demands and needs across areas. Experian developed a tool, Mosaic, to segment residents based on characteristics such as property and lifestyle. The graph in Figure 1 shows one of the ways Mosaic groups individuals. It groups individuals into 15 summary groups and 66 detailed types, which are given codes and names based on similar demographics, lifestyles and behaviours.*

By using GIS, I mapped the distribution of mosaic groups. For example, Figure 2 illustrates the distribution of the "I39 - Families on a Budget" segment, which includes families with children living in low-value social housing and making careful use of limited resources.

Finding the distribution of certain groups will support the targeted policy development.

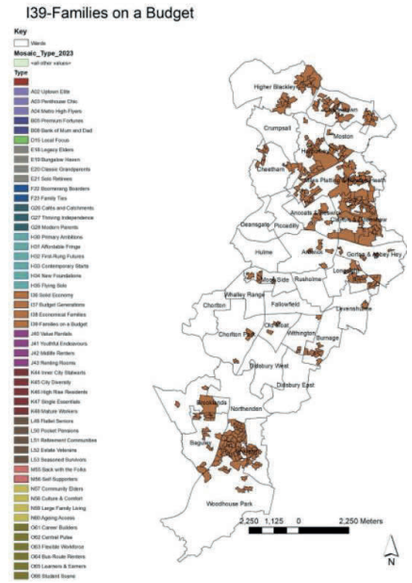


Figure 2- Spread of I39 in Manchester

For example, if the certain group has a high health risk, the following policies can be made to increase the healthcare resource allocation to this area.

Another task was verifying the validity and coherence across datasets. For instance, I compared income data from two datasets. Using Power Query, I analysed discrepancies in reported wages for the same postcode level, finding significant differences. Then, I quantify these disparities and identify regions or population groups with the inconsistencies.

Since these datasets are confidential, specific details are not presented here.

Key Skills Learnt

- *Data Analysis and Interpretation-The training provided taught me how to use ArcMap GIS and Excel effectively for data research. One memorable training session is how to identify reliable data sources and apply them to various research projects.*
- *Professional Environment Familiarisation- I participated in meetings with various teams and external stakeholders, both online and offline. This experience helped me to understand how ideas were exchanged and work progressed in a professional setting.*
- *Understanding of England's public sector- Daily interactions with government workers provided a valuable perspective on the UK government's operations. Also, I observed a full council meeting, gaining first hand insight into democratic decision-making processes.*
- *Improvement of Argument Skills- This internship helped me adopt a broader perspective by stepping outside my student mindset to understand the rationale behind government solutions and allowing me to approach social issues with greater comprehension and analytical depth.*

Figure 1-Experian's Mosaic Types

Department for Education: What is the main barrier to early years settings expanding provision?

DATA FELLOW AT EARLY YEARS ANALYSIS AND RESEARCH UNIT

Xuanzi Zhao | BA(Hons) Linguistics

Overview of the Data Fellowship

During my Data Fellowship at the Department for Education (DfE), I analysed Early Years Foundation Stage Profile (EYFSP) data to support policy recommendations, particularly in alignment with the new Labour government's

Opportunity Mission. My responsibilities included designing and implementing a survey template, conducting cost-benefit analyses, and collaborating with analysts to provide actionable insights on childcare capacity and children's performance in early years education.

This role allowed me to develop technical skills in SQL and Excel, and gain a deeper understanding of how data-driven insights can shape effective policy.

Data Analysis

I conducted analyses that offered quick insights into how children's characteristics influenced their performance at the end of their reception year, as assessed by the EYFSP.

Supporting the Opportunity Mission's goal of increasing the percentage of children achieving a Good Level of Development, I examined published findings and underlying data, using SQL to quantify ambitious yet achievable metrics. For instance, I benchmarked the current attainment of girls with no disadvantage factors born in September to set realistic goals.

My approach combined qualitative and quantitative methods, involving a thorough review of academic literature from sources like the Education Endowment Foundation.

This experience not only improved my data interpretation and cost-benefit analysis skills but also deepened my ability to apply an evidence-based framework to policy recommendations effective policy.

Findings

Collaborating with colleagues, I produced a detailed report on childcare capacity and children's performance at the end of the reception year, supported by visualizations and strategic policy recommendations. Our findings created a foundation for future policy direction, supporting the Department's efforts to reach ambitious child development targets under the Opportunity Mission.

By setting clear metrics and evidence-based goals, we outlined strategies that could be pursued at varying budget levels, thereby offering data-driven guidance for strategic decision-making.

Key Skills Learnt

Communication: I learned to communicate effectively within a professional setting, including through presentations focused on conveying data insights accurately to civil servants.

Time Management: Balancing cost-benefit analysis with other responsibilities, I honed my time management skills by setting clear plans and dividing tasks into manageable steps.

Creativity: In self-guided analysis, my line manager encouraged me to explore various possibilities without predefined "right" answers, fostering creativity in my approach.

Adaptability: Working across different projects required me to adapt quickly to new tasks and collaborate with diverse teams.

This fellowship has been a transformative experience, enhancing my technical skills, broadening my understanding of policy work, and shaping my career ambitions. Working in a collaborative environment, I gained practical skills in data analysis, survey design, and research application, contributing to initiatives that support early childhood development and align with the Opportunity Mission.

As an international student and non-native English speaker, I embraced the challenge of meeting the role's demands and worked hard to strengthen my skills and knowledge, and this experience was pivotal for my professional growth, providing insight into the educational field and reinforcing my commitment to impactful, data-driven work.

I am grateful to the Data Fellowship program for championing equal opportunities for international students, allowing us to make meaningful contributions within a professional setting.