Frequent Service Users

Data Analytics Lab

May 2023

This project aims to discover members of the public with whom we have the most interactions across the organisation as a whole – our most frequent service users.

Table of Contents

1 Introduction

The aim of this project is to discover members of the public with whom we have most interactions across the organisation as a whole – our 'frequent service users'. The output will be the development of a dataset and Business Insight (Qlik) dashboard capable of displaying individuals' interactions with West Midlands Police (WMP), including phone contact, physical attendance, investigations of crimes, time spent in custody and missing person episodes. This presents an opportunity to identify those people whose circumstances are such that they are in regular need of a policing service and are more likely to experience some form of vulnerability.

The purpose of the analysis is to provide a holistic view of our most frequent service users so that we can identify members of the public who have high a volume of interactions with WMP, potentially across a number of departments. By combining data across a range of interaction types and processes the intention is to discover those people who may be missed by simply counting the number of times they call the police, or the number of times they are a victim. Where we are able to identify an intensive use of resources focused towards an individual or household, there may be an opportunity to better manage risk and vulnerability associated with these individuals and as a consequence manage responses more effectively.

This will help improve our service and facilitate the provision of appropriate multi-agency support. It is the remit of Local Policing Areas (LPAs) to work closely with communities, partner agencies and third sector organisations and the additional organisational intelligence generated by this tool will support their work.

2 Context

WMP already has tools available to identify 'repeat locations' and 'repeat callers.' Repeat locations are often places such as large supermarkets or hospitals where the density of footfall and nature of the activity results in a high number of calls for service. Similarly, repeat callers tend to be either partner agencies such as the ambulance service requiring our assistance; or people who are identified as having poor mental health and for whom support is required to address wider issues; or those whose situation makes them vulnerable and in need of police assistance. Local policing teams work closely with repeat callers, whether they are large organisations or vulnerable individuals and regularly review the proactive management plans (PMPs) created to address such issues.

Neither of these metrics provide a true reflection of whether there are certain individuals whose circumstances mean they require an intensive use of the organisation's resources across a number of departments. Initially, the request was to understand all incoming and outgoing communication with service users in terms of phone calls, emails, attendance, investigation effort and time spent in custody. However, it is not deemed proportionate or practical to access outgoing phone and email records to understand the volume and nature of outgoing communication traffic, so this element of the request has not been scoped.

The initial exploratory data analysis (EDA) focused on two specific questions; how to define a frequent service user and how this should be measured. It was decided that data from across the Connect (investigations and custody) and ControlWorks (and Avaya for call handling) systems would be used to capture the interactions across a number of touch points. Frequency of interactions is measured using multiple count totals which are combined together into one metric which can be ranked to form the list of most frequent service users.

3 Intended activity resulting from the project

Using the dataset created during the project, a Business Insights (Qlik) dashboard has been created which presents individuals identified as the most frequent service users. The data going into this dashboard would be updated monthly and uses data from the previous 12 months.

The dashboard will primarily be for the use of local policing teams in alignment with the new local operating model that was introduced in April 2023. Just as local neighbourhood officers have a good understanding of the top repeat locations and repeat callers, the data in the dashboard will provide an evidence base about which individuals and households in their local community have cause to interact with WMP on multiple occasions and thus ensure support provision activity is focused on the most relevant individuals and households. The local policing teams are best placed to engage with partner agencies and third sector organisations to support any vulnerability that may be identified as a cause of the higher volume of interactions. Equally, neighbourhood officers will be able to provide feedback about organisational processes which are generating repeat demand (for example victims calling to find out about the progress of an investigation).

The output of the dashboard could be used in submissions to the Crown Prosecution Service as evidence of the high level of demand created by persistent offending committed by some individuals or against some victims, or as evidence for Domestic Violence Protection Notices (DVPN) or similar. This information could also be used in discussions with partner agencies to evidence the need for a multi-agency approach with some individuals or households.

See section 6 for details of the dashboard.

4 Ethical Considerations

The aim of this project is to identify those individuals whose circumstances are such that they require an intensive use of WMP resources. This extends to situations where individuals are the suspects, offenders or victims of crimes. This will offer opportunities to improve our service to vulnerable people in the face of limited resources. There is no intention to 'flag' frequent service users in any way which would lead them to receiving a reduced service level. Rather, the intention is to identify those vulnerable individuals so that we can explore better ways of responding to their needs and utilise the full suite of options available, including a multi-agency response such as the Right Care, Right Person approach developed by Humberside Police. It is acknowledged that there will be some who may continue to require an intensive level of assistance because their particular circumstances require it. This tool will ensure that local policing teams are aware of these individuals and their circumstances and to ensure the wider organisation is sighted so that safeguarding them remains a priority.

This project contains no predictive element about people's future behavior and it does not aim to find relationships between features such as where individuals live or their age. This tool merely combines data we have about individuals' interactions with WMP into one place and ranks them by volume to assist with assistance decisions / further discussions. There is no special category data used in the tool.

The dashboard will only be accessible to specified users due to the individual level data that it contains, all end users will have clear information available to them about the methods used to create the ranking of individuals to ensure the output is used in the intended way. It is anticipated that the dashboard will provide a better overview of all available data without the need for large quantities of manual work to identify the frequent service users.

5 Methods

This project requires the use of data from across multiple systems, which needs to be processed and joined before it can be inputted into the Business Insights dashboard. All data used in the project is from the previous 12 months from the day the data processing is run. It is expected to be updated either monthly or quarterly as the outputs are not likely to change daily, so a longer time period is required for updates to show changes in the output.

The data sources used in this project are:

DATASET	SOURCE	DETAILS
Incoming phone calls to the contact centre	Avaya	Call duration
Records of contact (roc) and incidents	ControlWorks	Incident details, resource allocation, caller details
Crimes and the resulting investigations	Connect	Investigation details, role in crime, outcomes
Custody records	Connect	Time duration in custody, outcome of custody
Missing persons data	Compact	Missing person reports

5.1 Stage 1: Data processing over all service users

Firstly, all individuals who have had interactions with WMP are ranked using a collection of overview metrics in order to get a list of the most frequent service users. Only those individuals that are the most frequent service users will be taken forward for the full data process to gather more detailed information about the full extent of their interactions with WMP. In order to do this, the overview metrics are based on data from two main systems which is extracted and summarised for each unique individual. At this stage of the process no details are extracted, only summarised totals and counts for each individual using a unique ID number.

5.1.1 Investigations and Custody (Connect)

The first main system used is Connect. All crime and custody records created in the past 12 months are linked back to any individual that had a role in them (e.g. offender, suspect, victim, person in custody), and for each unique ID, totals are calculated for:

1	**Suspect Total**	A count of the number of crime investigations where individual is the suspect.

2	**Offender Total**	A count of the number of crime investigations where individual is the offender.
3	**Victim Total**	A count of the number of crime investigations where individual is the victim.
4	**Custody Total**	A count of the number of custody records per individual, where applicable.
5	**Custody Duration**	Time duration spent in custody, in hours. Any durations above 200 hours are assigned the value 200. Any negative durations (assumed to be data entry errors) are assigned values of 0.

Each of the above totals has a different range of values, and the time duration is on a different scale to the count totals. To overcome any skewed results when combining totals, they are all normalized (by dividing each value by the mean of the top 1000 values for each category) and scaled to have values between 0 and 1.

Combining these five totals together gives an overall metric for investigations and custody that can be used for ranking individuals. In the final dashboard this metric is known as '**Investigations and Custody**' (Metric 1).

5.1.2 Calls for Service (ControlWorks)

In the same way that the Connect data was summarized into a metric that can be ranked over all individuals, the ControlWorks data was used to calculate the number of calls for service and the time spent responding to those calls. This data can be looked at from the perspective of the person making the call (**originator**), or the person that the call was made about (**subject**).

For each controlworks incident (or RoC), the information about the originator and person, along with the total duration of time between all resources that were assigned to the incident. Start time for each resource was taken to be assigned, deployed or arrival time (whichever came first) and end time for the duration was released time. The total duration for all resources for each incident was calculated, in hours.

For every incident in the above dataset, the information was linked twice to the dataset containing individuals. Firstly, linked against the originator of the call using the originator forename, surname, date of birth and PNC ID where available. Secondly, linked in the same way but to the 'subject' information from the call (who the call was about). As this was all reliant on manually inputted information from the ControlWorks incident logs, the data can only be linked to an individual when the record is accurate and complete. Not every ControlWorks incident can be linked to both the originator and the person information, and in some cases, there are no matches. Incidents that were pre-planned events or operations were excluded from the ControlWorks data.

The number of incidents per individual were counted, where they are the originator and / or the subject of the calls. The total number of resources used per individual was also summed, this gave four new totals:

6	**Originator Calls Total**	A count of the number of CW logs where the individual is the originator of the call.
7	**Originator Resource Duration**	The total duration (hours) of the resources that are linked to and dispatched due to the CW logs where the individual is the originator of the call.
8	**Subject Calls Total**	A count of the number of CW logs where the individual is the subject of the call.
9	**Subject Resource Duration**	The total duration (hours) of the resources that are linked to and dispatched due to the CW logs where the individual is the subject of the call.

These totals were also normalized and scaled to have values between 0 and 1 in the same way that the investigation and custody totals were. This allowed them to be combined with each having equal importance. Two metrics are calculated here, an **originator total** (Metric 2) and a **subject total** (Metric 3).

5.1.3 Ranking Individuals

The table below shows the summary of the nine totals that create the final four ranked metrics. From each of these metrics, a list of the top 100 individuals was extracted, giving a list of 400 individuals. Due to the likely overlap between individuals from each list, only distinct individual IDs were then taken forward for the full data extraction process. The final number of individuals displayed on the dashboard is therefore between 200 and 300.

Table 1: The four metrics calculated for each individual and which counts and durations they include. Metric 4 is an overall total that combines all 9 totals equally.

		Metric 1: **Investigation and Custody Total**	Metric 2: **Originator Total**	Metric 3: **Subject Total**	Metric 4: **Overall Total**
1	Suspect Count	✓			✓
2	Offender Count	✓			✓
3	Victim Count	✓			✓
4	Custody Count	✓			✓

5	Custody Duration	✓			✓
6	Originator Call Count		✓		✓
7	Originator Resource Duration		✓		✓
8	Subject Call Count			✓	✓
9	Subject Resource Duration			✓	✓

5.1.4 Main Process

For each of the individuals who were identified in the ranking process as frequent service users, a second data process is run to extract all relevant and more detailed information. This is summarized into the correct format to be displayed in the Business Insights dashboard. This is an iterative process carried out one by one for each individual.

1. For the selected individual, all versions of that person on our system are gathered (some people have the same ID, but different versions of spellings of their names). Lists of all possible forenames, surnames, DOB and PNC IDs associated with the individual are created, these lists are later used within other queries to search for the selected individual.
2. **Custody**: details of the number of custody records, duration of custody and file types of the associated cases are summarized for the selected individual.
3. **Investigations**: details of each investigation that the selected individual plays a role in, summarizing what role the individual played, the harm score of the crime, location of the incident and home address of the individual. This allows summarization of:
    a. Number of investigation locations at each address, and if the address is of someone they are related to.
    b. Number of investigations for each role in crime (offender / suspect / victim / witness / other).
    c. Outcome of each investigation (if there is no outcome, or it is a non-crime, the main group of the investigation is used instead)
    d. For investigations where the selected individual is the offender or suspect, establish the relationship of them to the victim.
    e. For investigations where the selected individual is the victim, establish the relationship of them to the offender or suspect.

f. For each person related to the selected individual, count the outcomes of the investigations.

4. **ControlWorks**: details of every ControlWorks incident (or RoC) log where the selected individual is either the originator of the call, or the subject of the call.

    a. Total number of CW logs associated with the selected individual.

    b. Using the originator phone number, establish relationship between caller and selected individual. Search through Connect data for person associated with each phone number and merge together where name and DOB match. If there is no record of the number on Connect from past interactions, but the name and DOB match with the name and DOB of a person from the relationship table (3.d/ 3.e) then assign relationship to number that way. Fill any unknown relationships with 'RELATIONSHIP UNKNOWN'. In some situations, the caller may be the selected individual.

    c. Handling time of the calls associated with the selected individual.

    d. Count of number of incidents created and number of RoCs.

    e. For all incidents, summary of incident types.

    f. Incident response grades.

    g. Location of incidents.

    h. Number of incidents that resources attended.

    i. Number of units from each resource group

    j. Hours spent at incidents from each resource group.

6 Dashboard

The dashboard consists of two pages – the main page and the details page, the contents and usage are described below.

6.1 Main Page

The main page of the dashboard shows a table containing the list of unique individuals that rank the highest for the metrics calculated over all individuals. The table displays their ranked scores for four metrics described in Table 1, alongside the 'harm caused' and 'harm received' for each individual. Harm caused is the total harm score (based on Cambridge Crime Harm Index (CCHI)) for all investigations that the individual is linked to where they are the suspect or offender. The harm received is the total harm score (CCHI) for all investigations where the individual is the victim. This data has not been processed in the same way as other harm scores (there is no time decay, etc.) and is provided merely as a means of gaining a picture of their vulnerability. This table also shows the LPA of the individual's home address, taken as the most common home address from all investigations they are linked to (there may be individuals who have no LPA in this table and this is due to them having an unknown home address on Connect).

A chart displays the breakdown of investigation location by LPA for each individual when selected as this may differ from their home address.

The table of individuals can be filtered or ordered in a number of ways. By filtering the table, the ranked metrics are not recalculated, but continue to show the overall rank out of all individuals. This is because there are not an even number of individuals per filter, and to ensure results are comparable across all individuals.

- By area: there is a filter button for each of the LPAs, and the table will be filtered to only show individuals with the selected LPA as their home address. This will allow local teams to easily view the people relevant to their local area. Local leadership teams will be able to decide on a manageable number for their neighbourhood teams to focus on and this will be reviewed through the normal tasking process.
- By theme: there are filters for:
    - DA (Domestic Abuse): investigations with domestic abuse keywords or crime flags
    - MH (Mental Health): investigations with mental health keywords or crime flags
    - CSE (Child Sexual Exploitation): investigations with child sexual exploitation or child sexual keywords or crime flags.
    - CA (Child Abuse): investigations with child abuse keywords or crime flags.
    - Missing persons: only individuals who have one or more compact reports from the last 12 months are displayed when this filter is selected.
- By priority crime type: for crimes that are deemed by Force and local tasking processes (such as violent crime or serious acquisitive crime).

6.2 Details Page

The second page of the dashboard displays further details for all of the information that is summarized on the main page. Only information about one individual at a time can be displayed. This page is split into sections for Investigations, Incidents, Custody Records and Arrests. There is also information displayed about resource time spent at incidents. Direct hyperlinks from the investigation and custody numbers to the record in the Connect system are included which take the user into the source system where they can view further details. This ensures only the minimum necessary information is displayed in the dashboard.

7 Development and Testing

To gain some feedback to implement into the development of the dashboard, it was released to new local performance teams from each LPA for testing purposes. They received an information and training input in March 2023 to enable them to understand the dashboard.

Whilst this is a new tool and the new operating model is still evolving, the early feedback has been positive:

- The majority of LPAs report that the list of names generated by the tool feel right to those with local knowledge of the area.
- The tool has highlighted a few names that were previously unknown to be frequent service users in some areas.
- The local performance teams are likely to make use of the tool on a monthly basis preparing for Local Tasking Delivery Boards and Safer Places meetings.

The main suggestions for amendments include how frequently the dashboard should be updated so that it aligns with the monthly meeting schedule; and which is the best home address to use to align people to the most appropriate LPA. The frequency of the update is being assessed by senior leaders to determine whether it should be monthly or quarterly.

The operational outputs being considered locally include:

- Creating a Proactive Management Plan (PMP) for an individual which ensures problem solving strategies are used and monitored by local neighbourhood teams.
- Sharing knowledge with partners at Safer Places meetings to identify any opportunities to collaborate.
- Ensuring there is adequate scrutiny around the progress of investigations into repeat offenders.
- Anti-social behavior (ASB) task and finish group with partners
- Referrals to other agencies depending on who is best placed to manage the need.

These operational activities are all standard neighbourhood policing tactics.