Department
of Health &
Social Care

Adult Social Care Client Level Data Dashboard

Methodology document

Version 6.2

Contents

Adult Social Care Client Level Data Dashboard 1
Contents 2
Introduction 3
Data processing 3
Fields derived in SQL 12
Metrics calculated within Power BI 14
SQL Code 15
Version control 15

Introduction

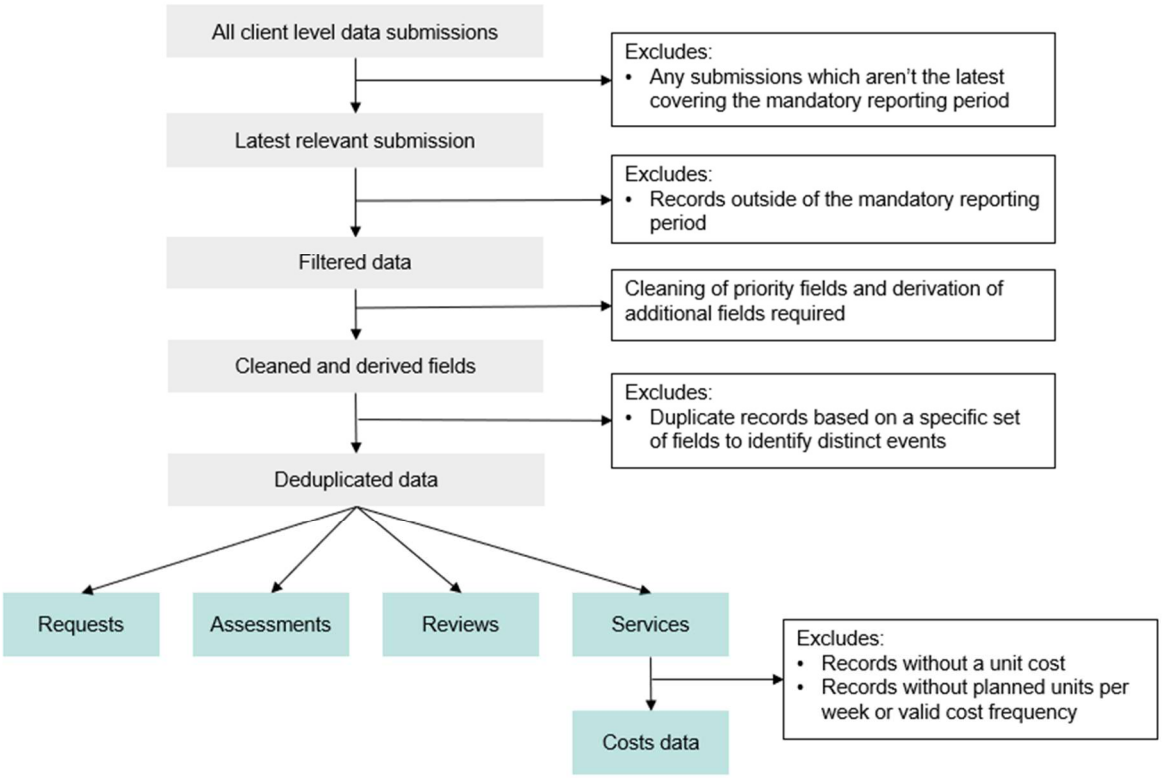
This document outlines the methodologies (logic and data processing) used to produce the data and metrics contained within the Adult Social Care Client Level Data dashboard for local authorities. It is split into 3 sections:

1. **Data processing** to generate SQL data tables that feed the dashboard
2. **Fields derived in SQL** using single records of data in the SQL data tables
3. **Metrics calculated within power BI** by performing calculations across a combination of fields and records from the data tables.

Data processing

Overview

This is an overview of the steps taken when processing Client Level Data to create the data tables which underpin the Power BI dashboard.



1. Filter to the relevant submission for each local authority

Data included in the dashboard is the latest CLD submission from each Local Authority covering the whole of the latest mandatory reporting period. The submission for each Local Authority is identified by selecting:

1. Files which cover the whole of the mandatory reporting period based on the Reporting Period Start Date and Reporting Period End Date stated in the submission. For example, for the first quarter's submission, files are selected with a minimum Reporting Period Start Date on or before 1st April 2023 and a maximum Reporting Period End Date on or after 31st March 2024.

2. If multiple files are returned in step one, then the file with the latest import date is selected.

To be included in the dashboard, the data must have been submitted as a single file covering the whole of the reporting period. Individual submissions covering consecutive months will therefore be excluded from this version of the dashboard. The data validation reports provided by AGEM contain a check for whether the submission meets the criteria for inclusion in the CLD dashboard, and this is fed back to Local Authorities.

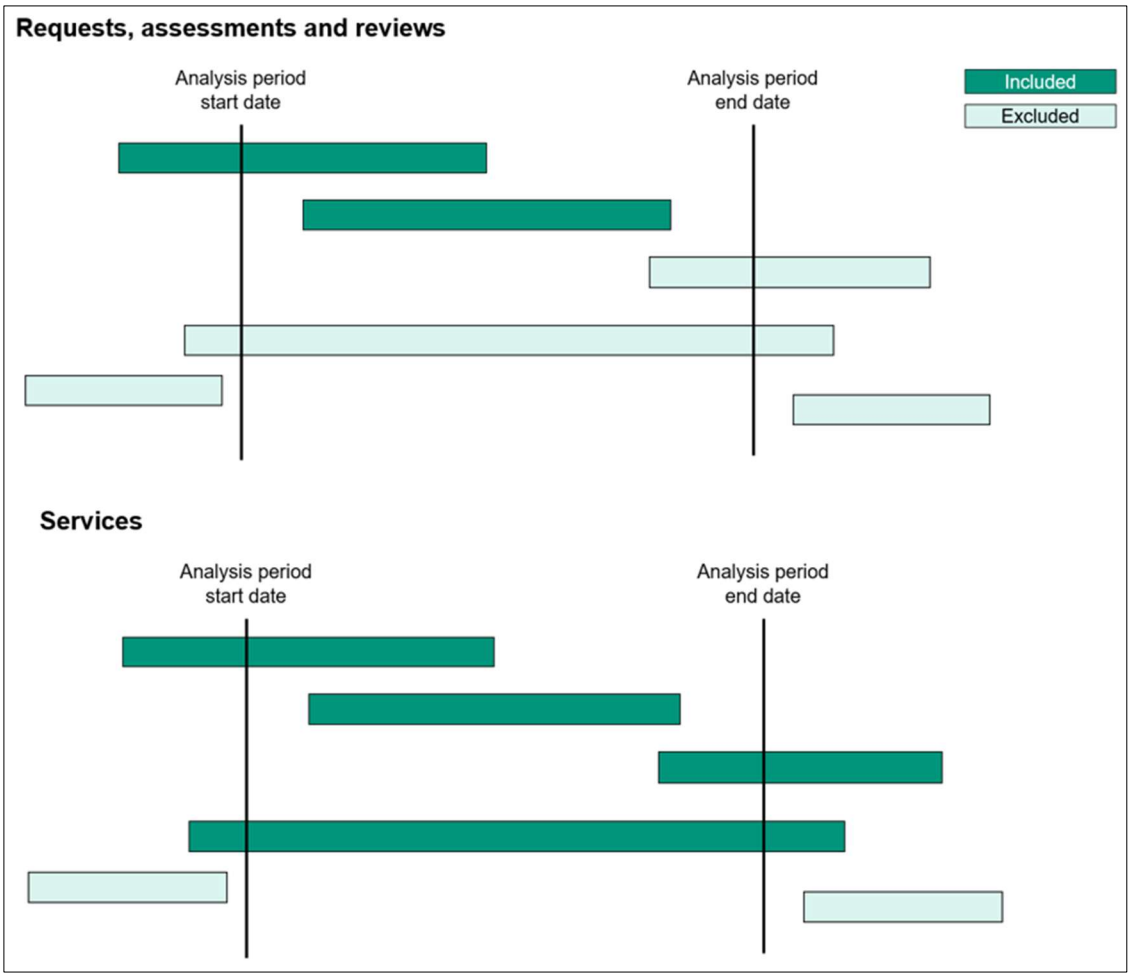
SQL scripts
Submissions are selected using the following procedures:
- REF_Submission_Reporting_Periods_table.sql – creates a reference table of one row per submission and their stated reporting period start and end dates.
- GetSubmissions – this procedure returns the most recent submission from each Local Authority that fully covers the period of interest (according to the stated reporting period start and end dates in the table above) and was submitted after the end of the reporting period and before the 'as of' date, if wanting to use a cut-off point for submissions.

2. Filter to the analysis period of interest

The dashboard shows events for the latest reporting period only. Inclusion criteria vary for the different event types, in line with the CLD guidance:
- Requests, Assessments and Reviews – include records where the Event Start Date is before or within the analysis period and Event End Date is within the analysis period.
- Services – Event Start Date must be before or within the analysis period and Event End Date must be within the analysis period or blank (open services).

Records with missing Event Type or Event Start Date are excluded at this stage and any requests, assessments and reviews with missing Event End Date are also excluded.

SQL scripts
Submissions are filtered to events within the period of interest using the FilterToEventsInPeriod procedure.



Accounting for date of death

Where a Date of Death is provided and it lies between the Event Start Date and Event End Date, or after the Event Start Date and where the Event End Date is missing, then the Date of Death is used as the Event End Date. This affects around 2% of records in the first quarter submissions. If the Date of Death is before the Reporting Period Start Date or before the Event Start Date, then this record is excluded. This affects less than 1% of records within the mandatory reporting period.

3. Cleaning and deriving new fields

For some priority fields, and fields impacted by the change from the release 1 to release 2 CLD specifications, data cleaning and/or mapping is carried out. For data submitted from October 2025 onwards, where local authorities have continued to use the previous Release 1 specification, the data is mapped to Release 2 where possible. This approach is necessary as all methods and scripts have been updated to support the Release 2 specification. No additional data rows are created, instead individual fields are amended as needed. Reference tables contain release 1 to release 2 mapping and are manually updated to map invalid entries to valid entries from the release 2 defined lists, where it is apparent what the entry should be.

At this stage derived fields are created, see the section on fields derived in SQL for more information.

SQL scripts
Some fields are cleaned and/or mapped, and additional fields are derived using the GetDerivedFields procedure.

4. Deduplicate to retain unique records only

We investigated ways of de-duplicating records to identify distinct records from multiple records. Following discussion with the local authority analytical working group and investigation of data quality across the first set of CLD submissions, we have used the fields listed in the table below to distinguish unique records within a single submission. More information on the methods for each event type is outlined below.

	Requests	Assessments	Services	Reviews
LA Code	✓	✓	✓	✓
Derived Person ID (NHS number unless missing then LA_ID)	✓	✓	✓	✓
Event Start Date	✓	✓	✓	✓
Event End Date	✓	✓		✓
Client Type	✓	✓	✓	✓
Request: Route of Access	✓			
Assessment Type		✓		
Service Type			✓	
Service Component			✓	
Delivery Mechanism			✓	
Unit Cost			✓	
Cost Frequency (Unit Type)			✓	
Planned units per week			✓	

SQL scripts
The Der_Unique_Event_Ref field is created in as part of the GetDerivedFields procedure. For requests, assessments and reviews this is derived using the fields listed above. For services, this field is based on distinct client type, event start date, service type and service component. It excludes fields expected to change between submissions (delivery mechanism, unit cost, planned units per week and cost frequency), however when using single submissions these are accounted for in the GetUniqueEvents script.
The GetUniqueEvents procedure deduplicates records using the Der_Unique_Event_Ref field. For requests, assessments and reviews this field alone is used. For services, this field is used in combination with delivery mechanism, unit cost, planned units per week and cost frequency when deduplicating to distinct events.

Reason for not using Event Reference Number
In the CLD guidance document, local authorities are asked to provide the Event Reference as a unique reference number for each event to help deduplicate records. Although intended for this purpose, in practice there are several issues that mean it does not work as intended:

- Event Reference is a voluntary field and around one in seven local authorities have not submitted it for all records.
- Some local authorities use the same Event Reference numbers by coincidence.
- In around 6% of records in the first quarter submissions, the Event Reference is identical, but values submitted for any of the other fields differ (e.g. Event Type, Event Start Date).
- In around 3% of records, the Event Reference differs, but values entered for all other fields are identical.
- Overall, use of the Event Reference number alongside the other variables appears to result in overcounting. It appears that some local authorities generate the Event Reference at the point of extracting the data from their case management system to create the CLD submission: it differs between each upload and does not correspond to a unique event.

Person identifiers
The anonymised person identifier used throughout the dashboard is based on the anonymised traced NHS number in the first instance, if this is missing then the local authority provided NHS number. If both NHS number fields are missing, the local authority person identifier is used. Please note, as of May 2024 this method was simplified and no longer includes an additional check for whether the local authority person identifier was previously associated with an NHS number elsewhere in the data since this step affected less than 0.01% of records.

Requests
The following fields are used to produce a distinct count of requests.

1. LA Code
2. NHS Number traced where populated otherwise LA Person ID
3. Event Start Date
4. Event End Date
5. Client Type
6. Request: Route of Access

If values submitted for all the above fields are identical, the data is ordered using the fields below and the first (top) record is retained:

Reporting Period End Date – descending (most recent at the top)
Reporting Period Start Date – descending (most recent at the top)
Event Outcome – as per the hierarchy in the CLD guidance
Derived unique record id – descending (most recent at the top)

Limitations:

- As with the SALT STS001, numbers of requests captured will be shaped by the local authority's arrangements for handling client contacts such as centralised call centres, outsourced arrangements etc. The method of recording client contacts locally will also influence what is shown on the dashboard.
- Where local authorities operate strengths-based approaches or offer drop-in assessments, initial contacts that are submitted solely as a record with Event Type = Assessment will not be counted as requests. To avoid this problem, we have suggested two approaches for local authorities when submitting data. We have recommended that these events should be submitted in the CLD return using Event Type = Request and that the Event Description field is used to indicate that these are initial conversations or assessments (Annex C of CLD guidance). Alternatively, some local authorities have submitted additional records for these events, one with the Event Type = Request and one with the Event Type = Assessment to ensure full information is submitted across all relevant fields. Either approach should ensure that initial contacts are counted as requests.

Assessments

The following fields are used to produce a distinct count of assessments.

1. LA Code
2. NHS Number traced where populated otherwise LA Person ID
3. Event Start Date
4. Event End Date
5. Client Type
6. Assessment Type

If values submitted for all the above fields are identical, the data is ordered using the fields below and the first (top) record is retained:

Reporting Period End Date – descending (most recent at the top)
Reporting Period Start Date – descending (most recent at the top)
Event Outcome – as per the hierarchy in the CLD guidance
Derived unique record id – descending (most recent at the top)

Because we have received different views from local authorities, the current page does not currently count initial conversations as short-term assessments. The CLD guidance advises that these events are submitted as requests under Event Type but notes that the Event Description field can be used to indicate that these are conversation records. We will review how to best reflect this data in future iterations of the dashboard following further conversations with local authorities operating strengths-based approaches.

Services
The following fields are used to produce a distinct count of services.

1. LA Code
2. NHS Number traced where populated otherwise LA Person ID
3. Event Start Date
4. Client Type
5. Service Type
6. Service Component
7. Delivery Mechanism
8. Unit Cost
9. Cost Frequency (Unit Type)
10. Planned Units per week

Event End Date is not included since open and ongoing services are submitted with this field left blank, and the field is only populated once the service is closed.

If values submitted for all the above fields are identical, the data is ordered using the fields below and the first (top) record is retained:

Reporting Period End Date – descending (most recent at the top)
Reporting Period Start Date – descending (most recent at the top)
Event Outcome – use the hierarchy in the CLD guidance
Derived unique record id – descending (most recent at the top)

Reviews
The following fields are used to produce a distinct count of reviews.

1. LA Code
2. NHS Number traced where populated otherwise LA Person ID
3. Event Start Date
4. Event End Date
5. Client Type

If values submitted for all the above fields are identical, the data is ordered using the fields below and the first (top) record is retained:

Reporting Period End Date – descending (most recent at the top)
Reporting Period Start Date – descending (most recent at the top)
Event Outcome – as per the hierarchy in the CLD guidance
Derived unique record id – descending (most recent at the top)

5. **Additional exclusions applied to costs data**

The costs dataset, which underpins the costs page on the dashboard, is derived from the services dataset with the following additional processing steps applied.

Only service events within the latest quarter are retained, and not the full mandatory reporting period. This is because we expect submissions to include the latest cost

information. For services spanning multiple quarters using the latest cost information may misrepresent services that have undergone uplifts. In longer time we intend to develop a methodology for using multiple submissions to analyse costs over time.

Planned units per week and Cost Frequency both provide information on the frequency of a service but there are instances where the two fields don't obviously align (e.g. Cost Frequency unit type = monthly but Planned units per week = 3). In general, Cost Frequency is trusted over the Planned units per week as Planned units per week is only required for services where the unit cost occurs more frequently than weekly e.g., hourly, daily, or per session.

To simplify calculation of metrics, average weekly units have been derived from Cost Frequency when the Cost Frequency is weekly or less frequent as follows:

Cost frequency (Unit Type)	Derived weekly units
Weekly	1
Fortnightly	1/2
4-weekly	1/4
Monthly	12/52
Quarterly	4/52
Annually	1/52

When the Cost Frequency is more frequent than weekly the absolute values are taken for Planned units per week. The absolute values are also taken for Unit Cost to deal with any negative values (~2% of service records). No further attempt is made to clean the data; therefore, any erroneous costs or outliers will be included in the dashboard.

Any records without a unit cost are removed. Any records without planned units per week (actual or derived) are also removed unless the cost frequency is 'one-off'. Because of higher rates of incompleteness of costs fields in the first set of submissions, these two exclusion steps result in the costs table containing approximately 25% fewer records than the services fact table.

6. Creating local authority level tables

Once a table of distinct records for each event type is created, this is aggregated up at local authority level. This is done by grouping by all the fields which are required to create the metrics and breakdowns in power BI. As this includes fields such as person ID and event start and end dates, the resulting aggregation is very minimal, and most of the data remains at record-level.

7. Data quality reporting

This page contains a completeness and validity check of each entry in each field, indicating whether each entry meets the expected data type and defined list of values. These checks align with the data quality reports provided by Arden & GEM back to local authorities (via MESH). The dashboard allows users to select whether the data field is applicable and mandatory for a given record, based on several key

fields (such as client type and event type). These filters are developed in line with the CLD guidance. For example, the data quality check of the Request: Route of Access field is only applicable and mandatory when a record is a request.

8. ASCOF Figures

The methodology for the ASCOF measures on the dashboard is published as a separate document on Athena and available on AGEM's website.

9. Publication statistics

The CLD dashboard contains metrics on long term support services and adult social care assessments, which are consistent with the metrics published in the Adult social care client level data, England: quarterly update. More information on the methods used to produce these statistics is available in the background quality and methodology document.

Fields derived in SQL

[Field_name]_Cleaned

For some priority fields, or fields which are impacted back the change from the release 1 to release 2 CLD specifications, data cleaning and mapping is carried out. Reference tables contain release 1 to release 2 mapping and are manually updated to map invalid entries to valid entries from the release 2 defined lists, where it is apparent what the entry should be. If the entry cannot be mapped due to data quality issues then it is overwritten to 'Invalid and not mapped'. These cleaned fields are used throughout the main dashboard pages, publication metrics and ASCOF metrics. The reference tables cannot be shared with local authorities, as they contain individual data sourced from local authorities CLD submissions.

Der_Latest_Age

This is a person's age at the latest point in an event. For all records where an end date is provided, this is the age at the Event End Date: for services where the Event End Date is blank, the Reporting Period End Date is used. As we do not have access to the date of birth, the Birth Year and Birth Month fields are used and the first day of the month is used as a proxy for the date of birth.

Der_Age_band and Der_Working_Age_Band

These fields assign age bands depending on the value of the Der_Latest_Age (derived above).

Der_Person_ID

This field contains the Der_NHS_Number_Traced_Pseudo unless this is missing, then the Der_NHS_Number_Pseudo is used unless this is also missing, in which case the local authority Person ID is used.

Review_type

This field assigns two high level categories of 'planned' or 'unplanned' based on the Review Reason field.

Service_type_grouped

For purposes of filtering the data on the dashboard, this field assigns three high level categories based on the Service Type: 'Short Term', 'Long Term' and 'Carer Support'

Event_outcome_grouped

This field groups together all Event Outcomes which relate to no further action (NFA), for the purposes of plotting this on a chart.

Event_outcome_hierarchy

The Event Outcome hierarchy field contains the ranking of the event outcome in line with the latest CLD guidance.

Long_Term_Service_Flag

The LTS Flag is a field added to the request, assessment and review tables. It determines whether the individual has a long term service event open at the time of the request, assessment or review start date. Please note, this is a proxy definition for a new client and does not align with the definitions used in the Activity statistics or the ASCOF measures.

Costs fields

Weeks_of_service – the number of weeks the service was open for within the reporting period. This uses the start and end dates of the event unless these fall outside of the reporting period, in which case the reporting period start and end dates are used. This step occurs after date of death has already been accounted for, as described earlier.

Cost_per_week – unit cost*planned_units_per_week, except when the cost frequency is 'one-off'. For unit costs that are less frequent than weekly, the derived weekly units are used e.g. for monthly service costing £500 per month, the cost per week will be £500*12/52 = £115.40 per week.

Total_cost_period – the total cost for the service within the reporting period. Where cost frequency is one-off then the unit cost is taken. For all other events this is Cost_per_week*Weeks_of_service.

Metrics calculated within Power BI

Number of requests received, assessments completed, and reviews completed over time

These measures are a count of the number of requests received, assessments completed, or reviews completed on a given day. For the number of requests received, this is based on the Event Start Date. For the number of assessments completed and reviews completed this is based on the Event End Date. These are aggregated up to weekly and monthly counts and displayed in the line chart visuals. For weekly reporting, any weeks at the start or end of the period which do not contain a full 7 days are not shown, to prevent the data being misrepresented.

7 day rolling average of requests received

This is only provided for requests received, and it calculates the average number of requests completed in the past 7 days, reported daily.

Number of active service users and carers

This measure counts the number of service users and carers with service events open and ongoing at a given point in time. This is determined by counting events with an Event Start Date on or before the given calendar date and an Event End Date which is either null, greater than or equal to the given calendar date. This is reported weekly by week ending.

Counts of individuals

The following measures are created by taking a distinct count of the der_person_id field for each of the event tables. For more information see Person Identifiers.

- Person_count_requests
- Person_count_assessments
- Person_count_services
- Person_count_reviews
- Person_count_costs

Mean cost in period per person

This is the total reported cost in the period divided by the total number of service users. This measure is an average across all service users and costs and therefore, does not reflect the intensity of a service. For example, one user may receive 6 weeks of service whilst the majority receive a couple of days.

SQL Code

The SQL scripts which produce the data tables underpinning the CLD dashboard are publicly available on GitHub here: DataS-DHSC/ASC-CLD-LA-Dashboard: The code used to produce the data tables underpinning the Client Level Data Dashboard for Local Authorities. (github.com).

Version control

Version number	Summary of changes
1.0	Initial version
1.1	Revised deduplication methodology for service events to include Unit Cost, Planned units per week and Cost Frequency (unit type) in the fields which identify unique events. Cost analysis revised to only include the latest quarter and not the full reporting period.
2.0	Addition of waiting times metrics. Simplified person identifier methodology to take the NHS number (traced taken over the local authority provided number) however if missing then the local authority person identifier is used. Removed the methodology for counting requests with 'conversation' in the event description as assessments, due to an issue in the code and more work is required to better understand how these events should be analysed. Updated age bands to align with those in the published CLD statistics and they are compatible with aggregation into the '18-64' and '65+' age bands.
3.0	Addition of 5 CLD-derived ASCOF measures.
3.1	Updated the timeseries graphs for requests, assessments, and reviews pages to include person- and event- level aggregation.
4.0	Updated to include more detail on the processing of single submissions. Removal of waiting times methodology whilst the metrics are being developed in response to feedback. Removal of ASCOF methodology as this is available in a separate methodology document.
5.0	Update to include methodology information for the two publication metrics (assessments and long term support statistics) added to the dashboard.
6.0	Updated to include information on how release 2 values are dealt with in the interim prior to the new specification becoming mandatory, and more information has been added on how the cleaned fields are derived.
6.1	Updated to include information on how release 1 values are dealt with since moving to Release 2 from October 2025.

6.2	Updated information on publication metrics to signpost to published background and methodology information.